



THE UNIVERSITY *of* EDINBURGH

## Edinburgh Research Explorer

### The transcription factor E2A drives neural differentiation in pluripotent cells

**Citation for published version:**

Rao, C, Malaguti, M, Mason, J & Lowell, S 2020, 'The transcription factor E2A drives neural differentiation in pluripotent cells', *Development*, vol. 147, dev.184093. <https://doi.org/10.1242/dev.184093>

**Digital Object Identifier (DOI):**

[10.1242/dev.184093](https://doi.org/10.1242/dev.184093)

**Link:**

[Link to publication record in Edinburgh Research Explorer](#)

**Document Version:**

Publisher's PDF, also known as Version of record

**Published In:**

Development

**General rights**

Copyright for the publications made accessible via the Edinburgh Research Explorer is retained by the author(s) and / or other copyright owners and it is a condition of accessing these publications that users recognise and abide by the legal requirements associated with these rights.

**Take down policy**

The University of Edinburgh has made every reasonable effort to ensure that Edinburgh Research Explorer content complies with UK legislation. If you believe that the public display of this file breaches copyright please contact [openaccess@ed.ac.uk](mailto:openaccess@ed.ac.uk) providing details, and we will remove access to the work immediately and investigate your claim.



# The transcription factor E2A drives neural differentiation in pluripotent cells

Chandrika Rao<sup>1</sup>, Mattias Malaguti<sup>1</sup>, John O. Mason<sup>2,3</sup> and Sally Lowell<sup>1,\*</sup>

## ABSTRACT

The intrinsic mechanisms that link extracellular signalling to the onset of neural differentiation are not well understood. In pluripotent mouse cells, BMP blocks entry into the neural lineage via transcriptional upregulation of inhibitor of differentiation (Id) factors. We have previously identified the major binding partner of Id proteins in pluripotent cells as the basic helix-loop-helix (bHLH) transcription factor (TF) E2A. Id1 can prevent E2A from forming heterodimers with bHLH TFs or from forming homodimers. Here, we show that overexpression of a forced E2A homodimer is sufficient to drive robust neural commitment in pluripotent cells, even under non-permissive conditions. Conversely, we find that E2A null cells display a defect in their neural differentiation capacity. E2A acts as an upstream activator of neural lineage genes, including *Sox1* and *Foxd4*, and as a repressor of Nodal signalling. Our results suggest a crucial role for E2A in establishing neural lineage commitment in pluripotent cells.

**KEY WORDS:** E2A, BHLH, Neural development, Pluripotent

## INTRODUCTION

Following the establishment of the primary germ layers during gastrulation, the embryonic neural plate is specified from the anterior ectoderm at approximately embryonic day (E)7.5 in a process known as neural induction (Tam and Zhou, 1996). Pluripotent embryonic stem cells (ESCs) recapitulate central features of this process when differentiated in culture (Ying et al., 2003b), providing a useful system in which to study the mechanisms guiding these initial cell fate decisions during development. Although the key extracellular signalling pathways that inhibit neural lineage commitment have long been established, less progress has been made in identifying the downstream effectors of these pathways. In particular, inhibition of BMP signalling is crucial for the establishment of the neuroectoderm (Harland, 2000; Di-Gregorio et al., 2007). The finding that BMP signalling inhibits neural differentiation via transcriptional upregulation of Id1 (Ying et al., 2003a) provides compelling evidence that Id proteins block

the activity of a factor that would otherwise trigger the onset of neural commitment.

Id proteins lack a DNA-binding domain and function primarily as dominant-negative inhibitors of basic helix-loop-helix (bHLH) transcription factors (TFs), binding to and preventing them from forming functional dimers (Norton, 2000). It has previously been reported that Id1 indirectly inhibits the activity of the bHLH TF *Tcf15*. However, *Tcf15* appears to play a role in general priming for differentiation by enabling morphological changes, and does not have a specific instructive role in neural commitment (Davies et al., 2013; Lin et al., 2019 preprint). We have also found that E-cadherin (*Cdh1*) acts downstream of BMP to help suppress neural commitment (Malaguti et al., 2013) but it is not known if or how Id1 is mechanistically linked with this process. Id1 is also reported to block the activity of the epigenetic regulator *Zrf1*, preventing derepression of neural genes in ESCs (Aloia et al., 2015). *Zrf1* overexpression alone, however, is not sufficient to drive expression of these genes, suggesting a requirement for additional factors to initiate neural differentiation in ESCs.

We have previously identified the alternatively spliced E2A gene products E47 and E12 as the main binding partners of Id1 in ESCs (Davies et al., 2013). E2A (also known as *Tcf3* – not to be confused with *Tcf7L1*, which is also commonly known as *Tcf3*) belongs to the E-protein family of bHLH TFs, which also includes HEB (*Tcf12*) and E2-2 (*Tcf4*). E2A is able to regulate the transcription of its target genes either by homodimerisation or by heterodimerisation with class II bHLH TFs, such as the proneural factors *Ascl1* and *neurogenin1/2* (Murre et al., 1989). Although E2A-bHLH heterodimers are well-established regulators of several fate determination processes, including neuronal subtype specification (Imayoshi and Kageyama, 2014), E2A homodimers have only been identified to function in the context of B-cell development (Shen and Kadesch, 1995), and it is not currently known whether this homodimer could also operate to control cell fate in other contexts.

E2A knockout mouse models have thus far failed to identify any overt gastrulation defects, with a failure of B-cell specification being the only major phenotype described to date (Bain et al., 1994; Zhuang et al., 1994). More recent analysis of these models, however, has noted that knockout mice have a significantly reduced brain size compared with their wild-type counterparts (Ravanpay and Olson, 2008), suggesting that a more in-depth investigation into the role of E2A during the earlier stages of development might be required to uncover subtle neural differentiation defects. In *Xenopus* embryos, loss of E2A has been associated with the inhibition of gastrulation (Yoon et al., 2011). Additionally, E2A and HEB have been shown to be co-factors of the Nodal signalling pathway, both in human ESCs and in *Xenopus* (Yoon et al., 2011), with E2A playing a dual role to directly repress the Nodal target gene *lefty* during mesendoderm specification in *Xenopus*, while also driving the expression of dorsal cell fate genes (Wills and Baker, 2015).

<sup>1</sup>MRC Centre for Regenerative Medicine, Institute for Stem Cell Research, School of Biological Sciences, University of Edinburgh, 5 Little France Drive, Edinburgh EH16 4UU, UK. <sup>2</sup>Centre for Discovery Brain Sciences, University of Edinburgh, 15 George Square, Edinburgh EH8 9XD, UK. <sup>3</sup>Simons Initiative for the Developing Brain, Hugh Robson Building, George Square, Edinburgh EH8 9XD, UK.

\*Author for correspondence (sally.lowell@ed.ac.uk)

© C.R., 0000-0002-5772-6772; M.M., 0000-0003-4392-7183; J.O.M., 0000-0002-0489-2400; S.L., 0000-0002-4018-9480

This is an Open Access article distributed under the terms of the Creative Commons Attribution License (<https://creativecommons.org/licenses/by/4.0>), which permits unrestricted use, distribution and reproduction in any medium provided that the original work is properly attributed.

Handling Editor: François Guillemot  
Received 21 August 2019; Accepted 26 May 2020

In this study, we investigate a potential role for E2A homodimers in neural fate commitment. We first set out to characterise the expression of E2A during early neural differentiation by the generation of an endogenously tagged E2A-V5 ESC line. Using a gain-of-function approach, we find that overexpression of a forced E2A homodimer, but not monomer, is sufficient to drive the neural commitment of ESCs, even in the presence of serum, providing novel mechanistic insight into the molecular events unfolding downstream of *Id1* during neural commitment. CRISPR/Cas9 targeting of E2A and HEB loci to generate single and double E-protein knockout ESC lines, additionally reveals that E-protein deficiency compromises the neural differentiation ability of ESCs. RNA-sequencing (RNA-seq) analysis further confirmed that E2A is positioned upstream of the expression of several neural lineage-associated genes, including *Sox1* and *Foxd4*, and might additionally play a role in suppressing Nodal signalling during neural differentiation. We therefore propose that the E2A homodimer is a key intrinsic regulator of neural fate commitment in pluripotent ESCs.

## RESULTS

### E2A is expressed heterogeneously in pluripotent ESCs and throughout neural differentiation

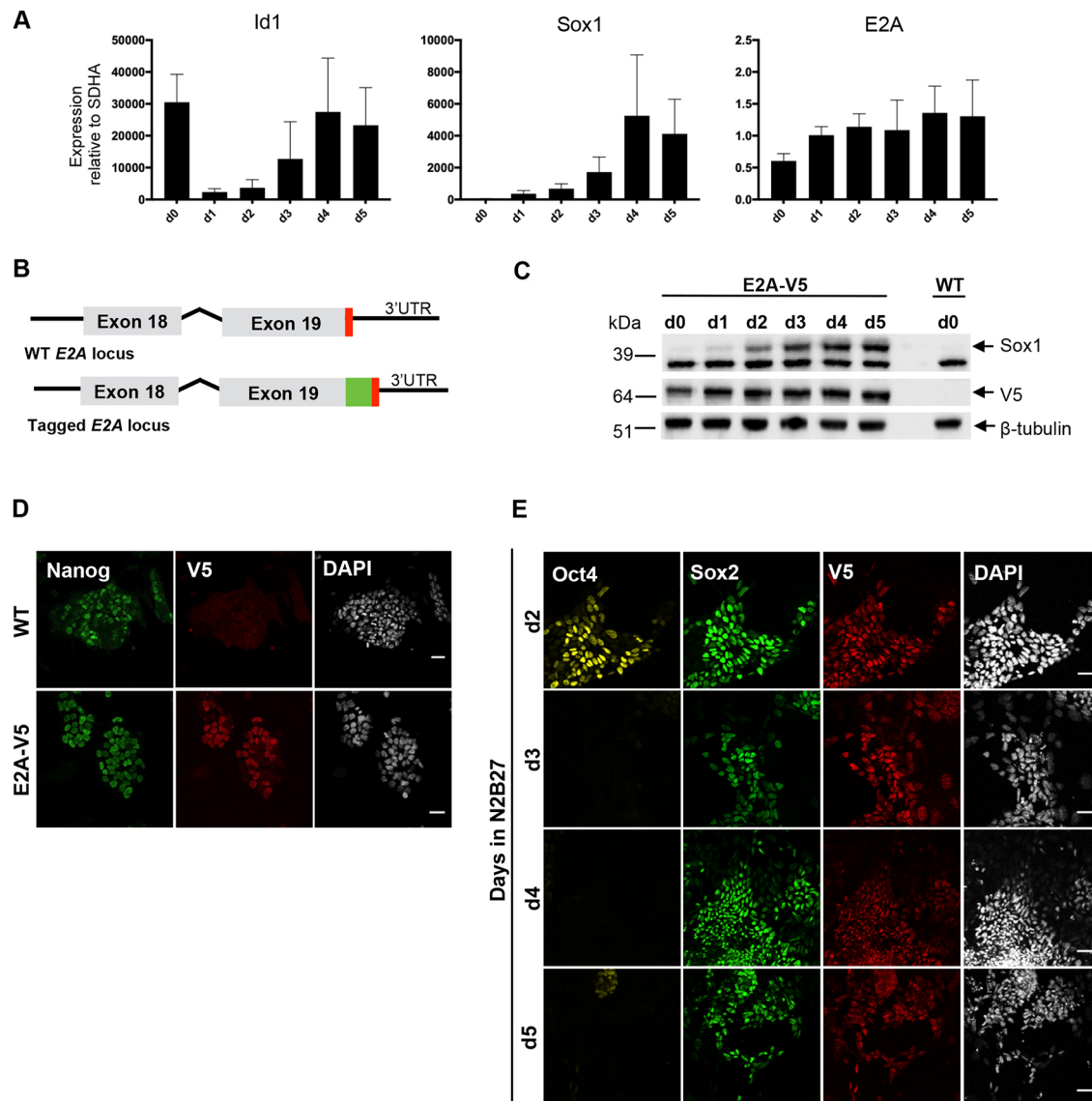
To characterise the expression of E2A during early neural differentiation we first examined the temporal expression of *E2A* mRNA in ESCs, plated under standard neural monolayer conditions (Ying et al., 2003b), by qRT-PCR. *Sox1* is the earliest specific marker of the neuroectoderm in mice (Pevny et al., 1998) and is therefore used to follow neural fate acquisition in ESCs. In line with previously published data (Ying et al., 2003b; Aiba et al., 2009), we observed that expression of the negative regulator of E2A, *Id1*, is rapidly downregulated at the onset of differentiation (Fig. 1A). *E2A* expression, however, remains fairly constant during this initial period. As E2A is regulated by *Id1* at the protein level, rather than the transcriptional level, we generated an endogenously tagged ESC line (Fig. 1B) using CRISPR/Cas9 targeting to follow the expression of E2A protein during differentiation. Based on a strategy previously developed to tag neural stem cells with high efficiency (Dewari et al., 2018), guide RNA (sgRNA) was designed to cut proximally to the stop codon in the 3' UTR of *E2A* (Fig. S1A), and was co-transfected into wild-type ESCs with recombinant Cas9 protein (rCas9) and a single-stranded donor DNA template (ssODN) encoding the V5 tag flanked by homology arms (Fig. S1B). Clonal lines were isolated from the bulk population and individual clones were subsequently screened by PCR genotyping (Fig. S1C), and analysed by immunostaining (Fig. 1D) for the V5 fusion protein. Sanger sequencing confirmed the error-free insertion of V5 into the C-terminus of the *E2A* locus by homology-directed repair (Fig. S1D).

We monitored the expression of the endogenously tagged E2A protein during neural differentiation using an antibody directed against the V5 tag for western blot analysis. This showed that E2A protein is expressed stably across the timecourse, with a slight increase in expression as cells exit pluripotency at day 1 (Fig. 1C), mirroring the pattern of *E2A* mRNA expression (Fig. 1A). Immunostaining of E2A-V5 cells revealed that E2A expression is heterogeneous in pluripotent ESCs on the single cell level in leukaemia inhibitory factor (LIF)/serum culture (Fig. 1D). When E2A-V5 ESCs are plated under neural differentiation conditions, E2A expression remains high and retains a heterogeneous expression pattern as cells exit the pluripotent state ( $\text{Oct4}^+\text{Sox2}^-$ ), and progressively commit to the neural lineage ( $\text{Oct4}^-\text{Sox2}^+$ ) (Fig. 1E).

### E2A homodimers promote neural fate commitment under self-renewal conditions

Having observed that E2A expression is dynamic during the early stages of differentiation, we next wanted to address whether E2A could play an instructive role during this cell fate transition. To determine whether E2A is sufficient to promote neural differentiation, we generated doxycycline-inducible ESC lines to overexpress either monomeric E2A or a forced E2A homodimer in which two E2A sequences are tethered by a flexible amino acid linker (Fig. 2A). This forced dimer strategy not only renders E2A more resistant to inhibition by *Id* but also favours the formation of E2A homodimers due to the physical proximity of the two molecules (Neuhold and Wold, 1993). FLAG-tagged E47 monomer or forced homodimer constructs were placed under the control of a tetracycline-response element and introduced into ESCs containing an inducible cassette exchange locus upstream of the HPRT gene (Iacovino et al., 2013). When expression of the two E2A constructs was induced by doxycycline (dox) in LIF/serum culture, conditions that are inhibitory for neural differentiation, we found that overexpression of the forced E2A homodimer, but not the monomer, elicited a robust upregulation of *Sox1* (Fig. 2B,D; Fig. S2). Coincident with the peak of *Sox1* expression on day 2, we also observed that cells overexpressing the homodimer lost the domed colony morphology typical of ESCs, and instead spread out to cover the dish. We evaluated that 34% of cells were  $\text{Sox1}^+\text{Oct4}^-$  after 4 days of culture (Fig. 2C), which was indicative of neural commitment. qRT-PCR analysis revealed that forced expression of E2A homodimers also enhanced the expression of the neural marker N-cadherin (*Cdh2*), downregulated *Cdh1* expression, and caused a transient upregulation of the epiblast marker *Fgf5* (Fig. 2D). Overexpression of both the monomeric and forced homodimer forms of E2A also appeared to stimulate a feedback response, causing upregulation of the negative regulator of E2A, *Id1*, as has been observed previously (Bhattacharya and Baker, 2011).

It has been reported that E2A homodimers are stabilised by an intermolecular disulphide bond formed between cysteine residues located in helix one of the bHLH domain (Benezra, 1994) (Fig. S3A). To confirm that it was specifically E2A homodimers, and not E2A-bHLH heterodimers, that are driving neural commitment in ESCs, we generated E2A forced dimer cell lines in which cysteine-570 is mutated to alanine in both E2A sequences (Fig. S3B) in an attempt to disrupt the formation of this covalent bond and therefore destabilise homodimer formation. These mutants are reported to retain heterodimerisation ability (Benezra, 1994). Two such mutant lines, C570A-1 and C570A-9, were stimulated with dox in LIF/serum for 2 days. qRT-PCR analysis of both mutant and control inducible lines showed that although C570A mutants were still able to upregulate *Id1*, they were no longer able to upregulate expression of *Sox1* (Fig. 2E), suggesting that this single amino acid substitution did, indeed, ablate the ability of E2A to drive differentiation. Western blot analysis of monomeric and forced E2A homodimer constructs demonstrated that expression of the forced homodimer was considerably lower than that of the monomer (Fig. S3C), ruling out the possibility that enhanced activity of the forced homodimer is explained simply by increased protein stability or the accumulation of E2A. C570A mutant lines appeared to express E2A at slightly lower levels than their wild-type homodimer counterparts. Taken together, these findings demonstrate that overexpression of E2A homodimers is sufficient to override the potent inhibitory effect of serum in non-permissive culture conditions, consistent with a role for E2A homodimers in initiating neural fate commitment.



**Fig. 1. Endogenous tagging reveals E2A is expressed heterogeneously in pluripotent cells.** (A) qRT-PCR analysis of pluripotent ESCs plated under standard neural differentiation conditions. Expression values are normalised to the housekeeping gene *Sdha*. Data are mean  $\pm$  s.d.  $n=3$  biological replicates. (B) Schematic of wild-type (WT) and tagged *E2A* locus. The V5 epitope tag was knocked-in at the 3' UTR of the endogenous *E2A* locus. The V5 tag is shown in green and the stop codon in red. (C) Western blot analysis of E2A-V5 cells during neural differentiation.  $\beta$ -Tubulin was analysed as a loading control. (D) Immunostaining of parental wild-type ESCs and epitope-tagged E2A-V5 ESCs. Cells co-stained for the V5 tag, the pluripotency marker Nanog and nuclear marker DAPI. (E) Immunostaining of E2A-V5 cells from days 2 to 5 of neural differentiation. Cells were stained for Oct4 and Sox2 to enable identification of cells committing to the neural lineage (Oct4<sup>+</sup>/Sox2<sup>+</sup>). Scale bars: 30  $\mu$ m.

### ***E2A*<sup>-/-</sup> and *E2A*<sup>-/-</sup>*HEB*<sup>-/-</sup> knockout ESCs display compromised neural lineage commitment ability**

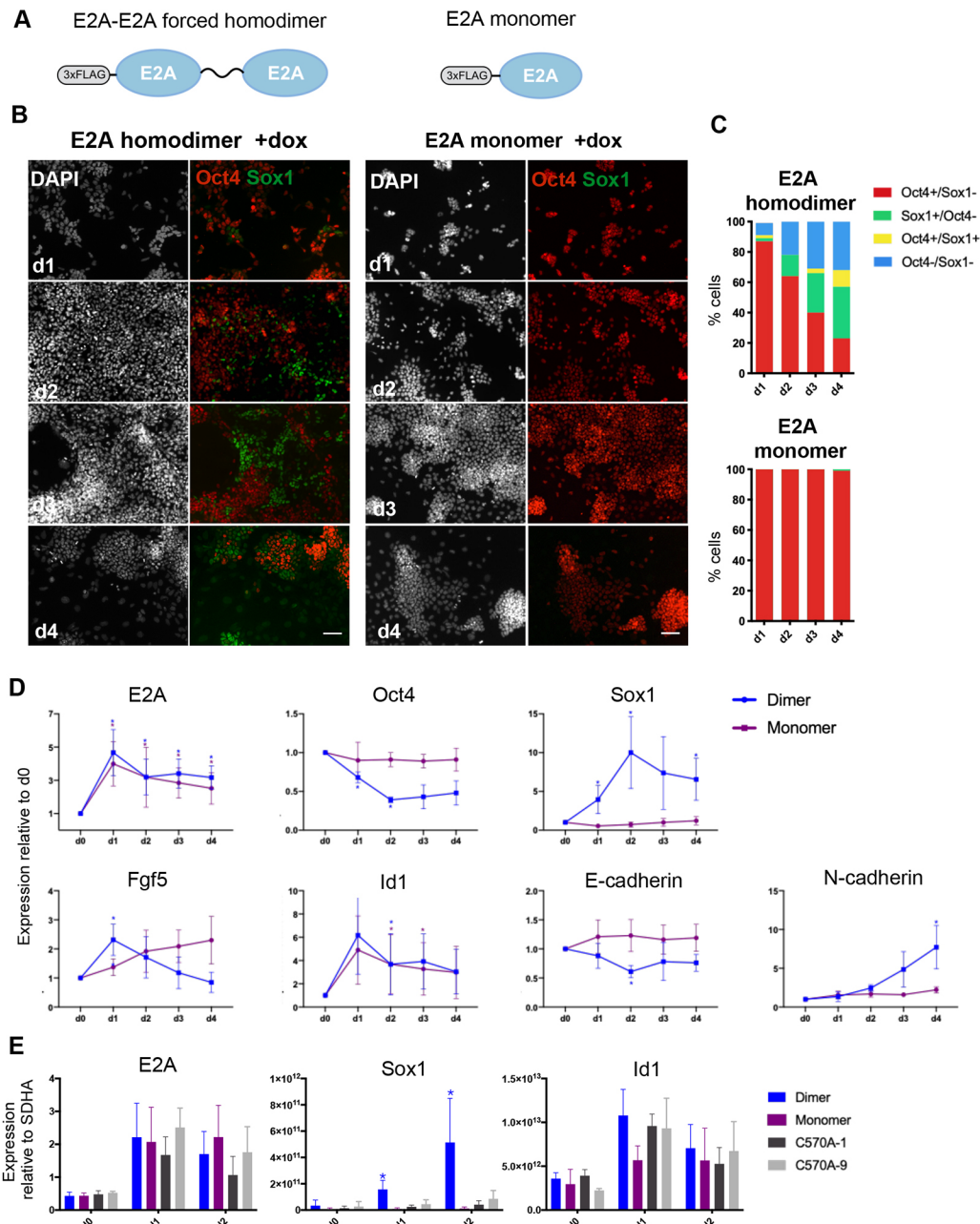
We next set out to evaluate whether *E2A* is necessary for neural differentiation. To generate *E2A* knockout ESCs, we used CRISPR/Cas9 to target exon 3 of the *E2A* gene locus, disrupting both *E12* and *E47* transcripts. Owing to the high degree of functional compensation observed between E-protein family members (Zhuang et al., 1998), which could mask potential differentiation phenotypes, we also generated *E2A*<sup>-/-</sup>*HEB*<sup>-/-</sup> double E-protein knockout ESC lines by using a similar strategy to target exon 9 of *HEB* in the *E2A*<sup>-/-</sup> cell line (Fig. S4A). Targeted lines were validated by western blot analysis and sequencing of the targeted loci (Fig. S4B-F).

To first determine whether E-protein knockout cells retain characteristic features of pluripotent stem cells, we assessed cell

morphology and gene expression in LIF/serum conditions. Colony morphology was indistinguishable between the parental *E2A*-V5 cell line and *E2A*<sup>-/-</sup> and *E2A*<sup>-/-</sup>*HEB*<sup>-/-</sup> cells, and immunostaining showed similar levels of expression of the pluripotency markers Oct4 and Nanog between knockout and parental cells (Fig. 3A), indicating normal self-renewal. qRT-PCR analysis of pluripotency and lineage markers in LIF/serum further confirmed that there were no differences in expression levels between parental and knockout cell lines (Fig. 3B).

We next examined the ability of the knockout cells to differentiate under standard neural monolayer conditions. qRT-PCR analyses showed that, compared with controls, *E2A*<sup>-/-</sup> and *E2A*<sup>-/-</sup>*HEB*<sup>-/-</sup> cells were unable to robustly upregulate markers of neural differentiation, including *Sox1*, *Foxd4* and *Pax3* (Fig. 3B; Fig. S5). Furthermore, although knockout cells were able to





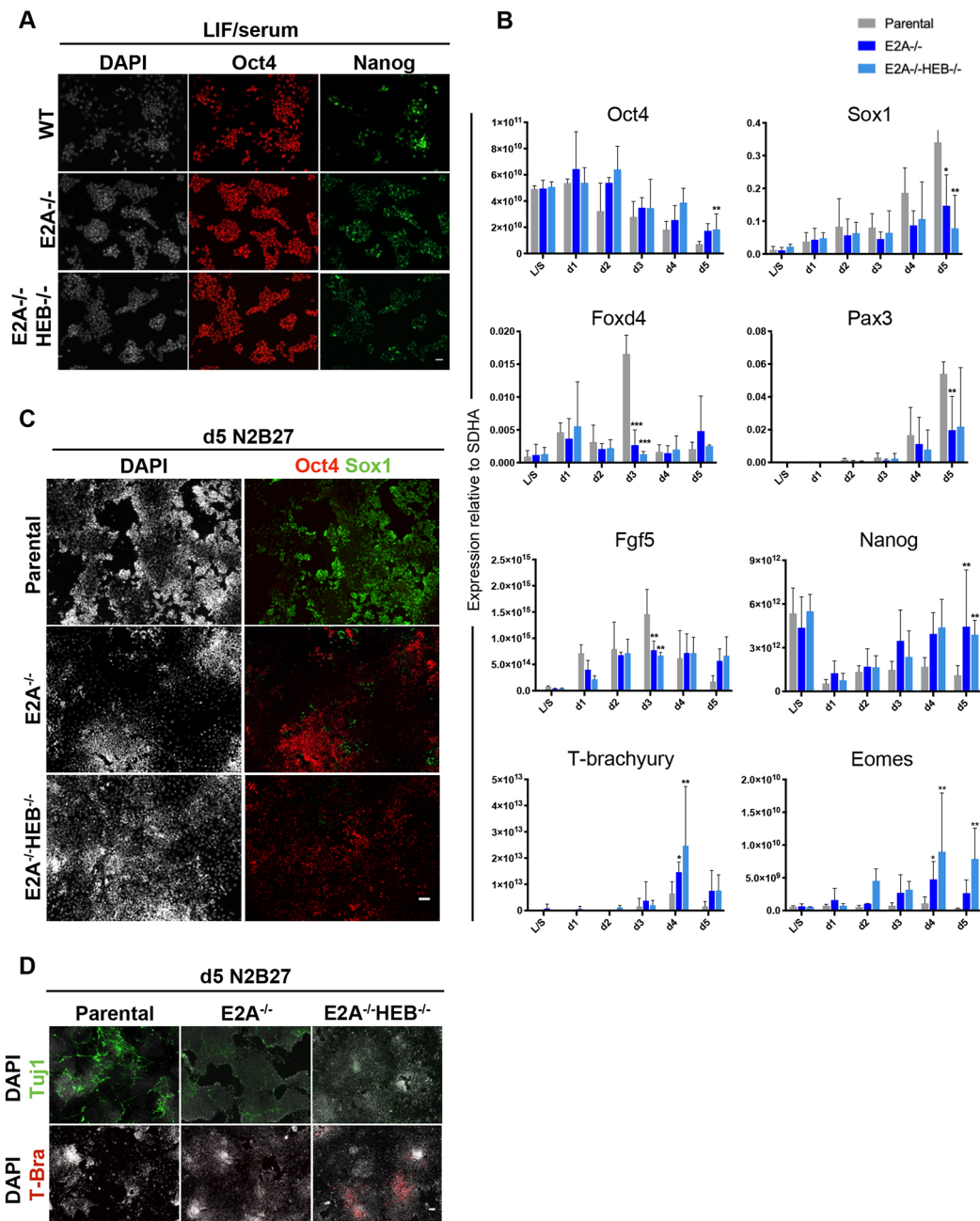
**Fig. 2. An E2A forced homodimer drives Sox1 expression under self-renewal conditions.** (A) Schematic representation of monomeric E2A and forced E2A homodimer constructs. (B) Immunostaining of inducible E2A forced homodimers and monomers cultured in LIF/serum+dox. Scale bars: 30  $\mu$ m. (C) Quantification of immunostaining for E2A monomers and homodimers cultured in LIF/serum+dox. A range of 500 to 5000 cells were manually scored for each timepoint. (D) qRT-PCR analysis of dox-inducible overexpression of E2A monomers and homodimers in LIF/serum. Expression is normalised to day 0 minus dox controls. (E) qRT-PCR analysis of two mutant forced homodimer clones (C570A-1 and C570A-9) induced in LIF/serum+dox. Expression values are normalised to the housekeeping gene *Sdha*. Data are mean $\pm$ s.d.,  $n=3$  biological replicates. \* $P<0.05$ , two-tailed paired Student's *t*-test.

downregulate expression of the naïve pluripotency marker *Nanog*, they maintained comparably high expression of the epiblast markers *Oct4* and *Fgf5*. Immunostaining of differentiation cultures for Sox1, Oct4 (Fig. 3C) and the marker of immature neurons, Tuj1 (Fig. 3D), further supported the gene expression data. Interestingly, we also observed that the disruption of E-protein expression resulted in an upregulation of the mesendodermal lineage markers, *T (brachyury)* and *Eomes* (Fig. 3B,D), as well as re-expression of *Nanog*, suggesting that these cells were beginning to adopt an identity more similar to the proximal epiblast, despite being cultured under neural differentiation conditions. Furthermore, we noted that both the neural differentiation

defect and the upregulation of mesendodermal markers were more pronounced when HEB was deleted in addition to E2A, indicating that HEB is able to functionally compensate for E2A, to a limited extent, in this context. Taken together with the overexpression data, these findings suggest that E2A is both sufficient and required for efficient neural differentiation of ESCs.

#### Identification of early E2A response genes

Having identified a novel role for E2A homodimers in driving the neural differentiation of ESCs, we next sought to identify the downstream targets of E2A in this process. We have previously



**Fig. 3.** *E2A*<sup>-/-</sup> and *E2A*<sup>-/-</sup>*HEB*<sup>-/-</sup> ESCs self-renew normally but display defects in early neural commitment. (A) Immunostaining of WT, *E2A*<sup>-/-</sup> and *E2A*<sup>-/-</sup>*HEB*<sup>-/-</sup> knockout ESCs stained for pluripotency markers in LIF/serum (L/S). (B) qRT-PCR analysis of parental and knockout cell lines in L/S and differentiated in N2B27 over 5 days. Expression values are normalised to the housekeeping gene *Sdha*. Data are mean±s.d., *n*=3 biological replicates. \**P*<0.05, \*\**P*<0.01 and \*\*\**P*<0.001, two-tailed paired Student's *t*-test. (C,D) Cells differentiated in N2B27 for 5 days stained for Oct4 and Sox1 (C), or the neuronal marker Tuj1 and the mesendodermal marker T-brachyury (T-bra). Scale bars: 30 µm.

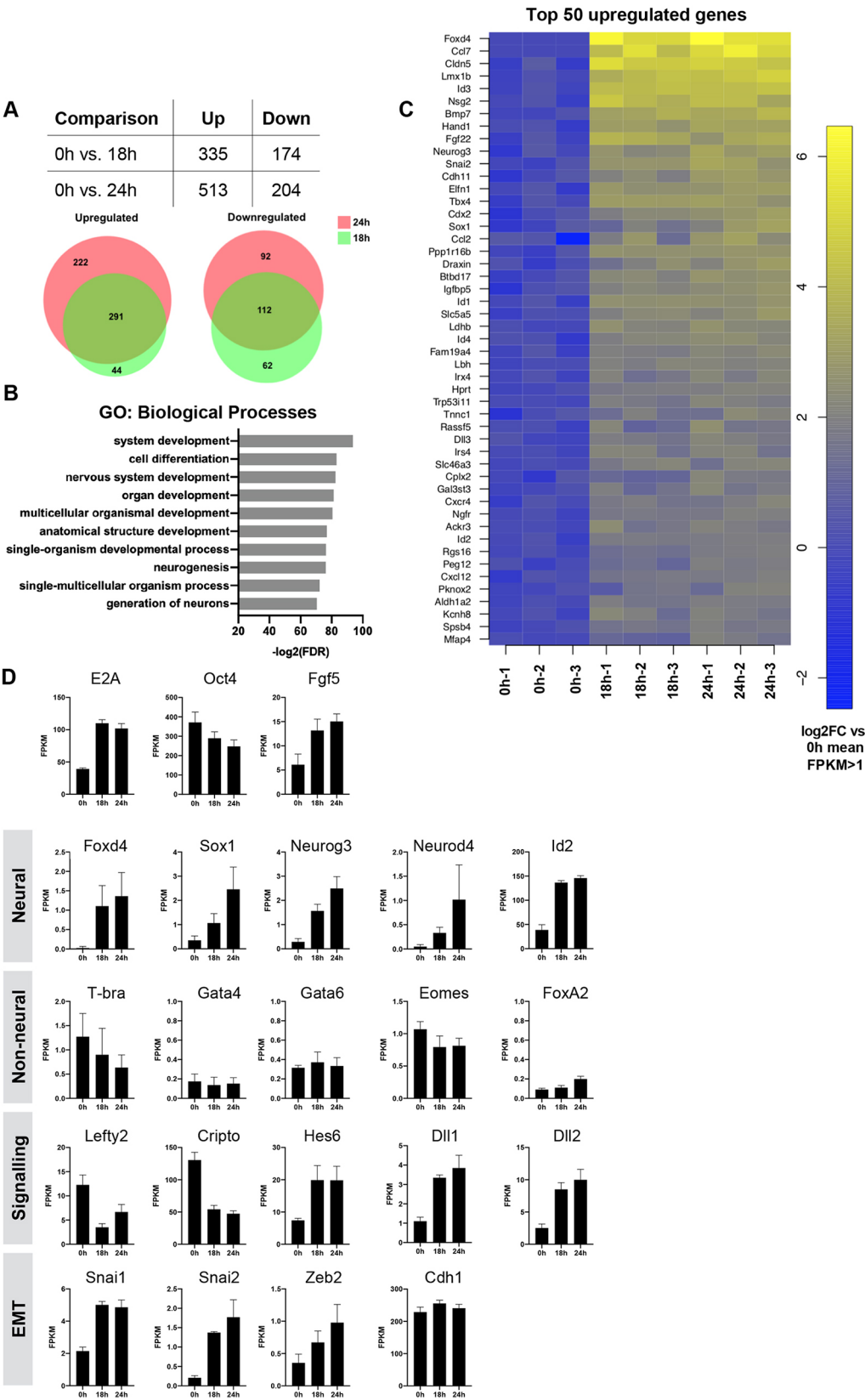
reported that BMP blocks neural differentiation by maintaining E-cadherin (Malaguti et al., 2013), and others have reported that E2A is a direct transcriptional repressor of E-cadherin (Pérez-Moreno et al., 2001). Combined with our earlier observation that forced homodimer expression causes a downregulation of E-cadherin mRNA expression (Fig. 2B), we hypothesised that E2A homodimers may also repress E-cadherin protein expression to enable neural differentiation. However, contrary to this hypothesis, we detected no significant downregulation of E-cadherin protein in response to the induction of E2A homodimers before the expression of Sox1 (Fig. S6A). Our findings do not exclude the possibility that

E-cadherin is a transcriptional target of E2A in pluripotent cells but they do suggest that E-cadherin is unlikely to mediate the effects of E2A on neural differentiation.

To identify novel downstream targets and examine genome-wide changes in expression in response to the induction of E2A homodimers, an RNA-seq approach was chosen. One aim of this approach was to identify factors that mediate the ability of E2A to upregulate Sox1. Expression of *Sox1* transcript was robustly upregulated by 18 h compared with the corresponding no dox control, whereas expression of Sox1 protein was first detected at 24 h (Fig. S6B,C). Based on these observations we performed

RNA-seq analysis at 0 h (no dox), 18 h and 24 h timepoints in triplicate to capture changes in genes acting upstream of Sox1. We found that of the 335 genes that were upregulated at 18 h compared

with 0 h, 291 remained significantly upregulated at 24 h, whereas 112 out of the 174 genes downregulated at 18 h remained downregulated at 24 h (Fig. 4A).



**Fig. 4. Global transcriptome profiling by RNA-seq reveals E2A homodimers specifically upregulate neural lineage markers.** (A) Differentially expressed genes in each comparison according to a minimum threshold of log2 fold-change of 2 and a maximum FDR of 0.05. Proportional Venn diagrams illustrate the overlap of differentially expressed genes between timepoints. (B) GO analysis of the top 50 upregulated genes at 24 h. (C) Heatmap of the top 50 genes upregulated at 24 h. (D) Mean FPKM (fragments per kilobase per million reads mapped) RNA-seq expression values for selected genes at 0 h, 18 h and 24 h. Data are mean±s.d., n=3 biological replicates.



## E2A homodimers specifically promote neural lineage commitment

Gene ontology (GO) analysis of genes upregulated at 24 h revealed an enrichment for terms associated with neural differentiation, including ‘nervous system development’, ‘neurogenesis’ and ‘generation of neurons’ (Fig. 4B). Further interrogation of the RNA-seq data highlighted an upregulation of neural lineage-associated genes, including *Foxd4* [log<sub>2</sub> fold-change (logFC) 5.7; false discovery rate (FDR) 0.0003], *Sox1* (logFC 2.8; FDR 0.003), *Neurog3* (logFC 3.1; FDR 7.18E-05) and *Neurod4* (logFC 4.4; FDR 0.02) (Fig. 4C,D). In contrast, early markers of non-neural lineages, such as *T*, *Eomes*, *Gata4*, *Gata6* and *Foxa2* were not upregulated, suggesting that E2A homodimers specifically drive neural lineage commitment in ESCs. We additionally observed an upregulation of EMT-related genes, including *Snail*, *Snai2* and *Zeb2*, but no detectable early downregulation of E-cadherin, which was in line with our previous qRT-PCR analyses. We also found an upregulation of Notch signalling components, including *Dll1*, *Dll3* and the FGF target gene *Hes6* (Koyano-Nakagawa et al., 2000; Murai et al., 2007), and downregulation of Nodal pathway genes, including the Nodal co-receptor *Cripto* (*Tdgfl*) and the Nodal target gene *Lefty2*. Taken together, these data suggest that E2A homodimers are able to dominantly induce a transcriptional programme associated with the early stages of neural differentiation, although it remains possible that additional factors might be required to enable cells to progress through the later stages of neuronal differentiation.

## E2A drives neural commitment by upregulating *Foxd4* and dampening Nodal signalling

The forkhead box TF *Foxd4* is a likely candidate for mediating the effect of E2A on neural differentiation. *Foxd4* was identified by RNA-seq as the most highly upregulated gene at 24 h following E2A homodimer induction (Fig. 4C), and is required for neural fate acquisition in mouse ESCs (Sherman et al., 2017). This is consistent with our observation that *Foxd4* is transiently upregulated at day 3 of neural differentiation, before overt *Sox1* expression, whereas *E2A*<sup>-/-</sup> and *E2A*<sup>-/-</sup>*HEB*<sup>-/-</sup> cells fail to upregulate *Foxd4* (Fig. 3B).

Alternatively, E2A could be modulating differentiation by dampening Nodal activity. The Nodal pathway genes *Cripto* and *Lefty2* were found to be downregulated in response to E2A induction (Fig. 4D). During gastrulation, the anterior epiblast fated for neuroectoderm is ‘silent’ for Nodal signalling (Peng et al., 2016), and inhibition of the pathway has been shown to promote neural differentiation in both mouse and human ESCs (Vallier et al., 2004; Watanabe et al., 2005; Camus et al., 2006). We found that although *Cripto* and *Lefty2* were rapidly downregulated and maintained at basal levels during the differentiation of control cells, they were progressively upregulated in E-protein knockout lines, with a more dramatic effect evident in *E2A*<sup>-/-</sup>*HEB*<sup>-/-</sup> cells (Fig. 5A).

Based on these observations, we hypothesised that *Foxd4* and/or a dampening of Nodal signalling could explain how E2A can drive neural differentiation. To test this, we sought to determine whether ectopic expression of *Foxd4* or inhibition of Nodal signalling could rescue the neural differentiation defect observed in E-protein knockout cells.

We first generated *E2A*<sup>-/-</sup> ESC lines that ectopically express *Foxd4*. A plasmid encoding *Foxd4* cloned upstream of an internal ribosomal entry site (IRES)-puromycin resistance cassette, under the control of a PGK promoter, was assembled and transfected into *E2A*<sup>-/-</sup> cells. *Foxd4* expression was assessed in undifferentiated ESCs by qRT-PCR (Fig. S7A) and three clonal lines were chosen for subsequent analyses (clones 3, 7 and 8). Ectopic expression of *Foxd4* was not sufficient to force neural differentiation of *E2A*<sup>-/-</sup>

ESCs cultured in LIF/serum, as assessed by *Sox1* expression (Fig. S7A). However, using quantitative immunostaining analysis, we observed that when cells were placed under differentiation conditions, ectopic expression of *Foxd4* was able to partially restore *Sox1* expression in knockout cells (Fig. 5B,C; Fig. S7B; Table S3) and to promote a large proliferation of the Oct4-expressing cells present in the culture (Fig. 5B,C).

We then sought to test whether inhibition of the Nodal pathway using SB431542 (SB43), which inhibits the activity of TGF-β receptors (Inman et al., 2002), could restore neural differentiation in E-protein knockout cells. We observed that SB43 treatment increased the proportion of *Sox1* expression in control cells, and resulted in a strong upregulation of *Sox1* in *E2A*<sup>-/-</sup> and *E2A*<sup>-/-</sup>*HEB*<sup>-/-</sup> lines. When applied to cells also ectopically expressing *Foxd4*, SB43 was able to effect a robust differentiation response (Fig. 5B,C; Fig. S7B,C). Taken together, these results suggest an additive effect of ectopic *Foxd4* expression and TGF-β pathway inhibition in restoring the differentiation capacity of *E2A*<sup>-/-</sup> cells. We conclude that E2A homodimers promote neural differentiation by upregulating *Foxd4* while dampening the activity of TGFβ ligands.

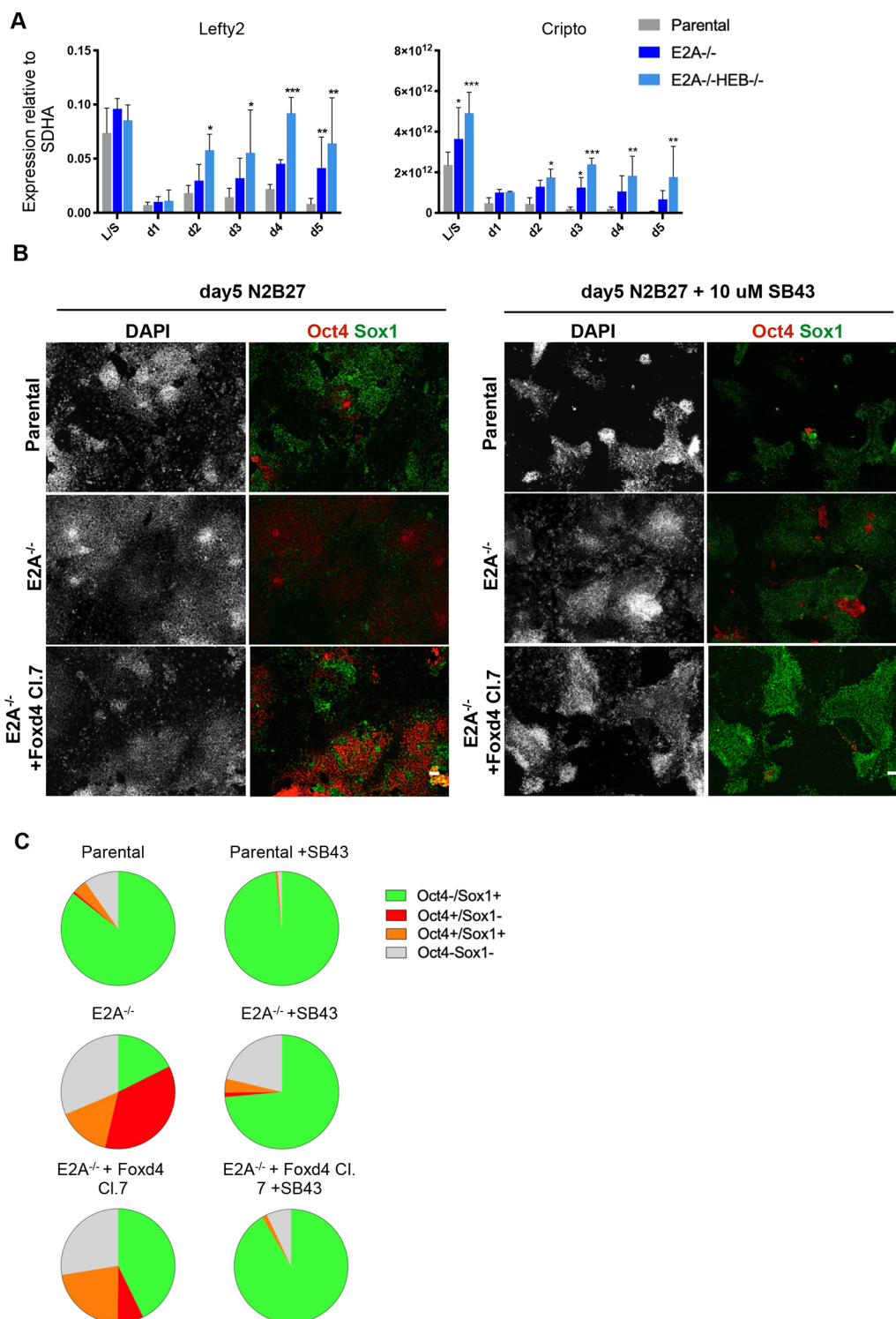
## DISCUSSION

The key extracellular signalling pathways that inhibit differentiation of ESCs are now well established (Haeghele et al., 2003; Smith et al., 2008; Ying et al., 2003a; Zhang et al., 2010). A considerable amount of progress has also been made in elucidating the network of TFs that control later aspects of mammalian neuronal specification and differentiation (Imayoshi and Kageyama, 2014). Much less is known, however, about the molecular processes that link these two events at the earliest stages of neural differentiation. Recent work has identified roles for TFs such as *Zfp521*, *Oct6* and *Zic1/2* as intrinsic regulators of early neural differentiation (Kamiya et al., 2011; Zhu et al., 2014; Sankar et al., 2016). However, although it has been demonstrated that *Oct6* is involved in the activation of several neural fate-promoting genes (Zhu et al., 2014), *Zfp521* and *Zic1/2* are more likely playing a role in consolidating, rather than initiating, neural fate (Kamiya et al., 2011; Iwafuchi-Doi et al., 2012). In this study, we sought to identify novel intrinsic regulators of neural fate commitment in ESCs, and to elucidate the molecular events unfolding downstream of the BMP signalling pathway during this process.

The HLH transcription factor *Id1* has previously been identified as a key effector of the BMP pathway, and its overexpression has been shown to block entry into the neural lineage and promote differentiation to alternative fates (Ying et al., 2003a,b; Malaguti et al., 2013). Although the function of *Id* proteins as dominant-negative inhibitors of bHLH transcription factors is well characterised in heterologous systems (Norton, 2000), it has not previously been investigated whether *Id1* could also be functioning via this classical mechanism during neural commitment. Taken together with the observation that E2A is the major binding partner of *Id1* in ESCs (Davies et al., 2013), we therefore hypothesised that *Id1* maintains pluripotency by sequestering E2A, thus preventing it from forming functional homodimers or heterodimers that could otherwise initiate a differentiation response.

Despite their widespread expression during embryogenesis (Pérez-Moreno et al., 2001), a definitive role for Class I bHLH TFs in early neural development has not been explored, with E-proteins often relegated to simply being obligate dimerisation partners for tissue-specific Class II bHLH factors. Single E-protein knockouts in mice have also failed to uncover any robust





**Fig. 5. Combined Foxd4 expression and Nodal signalling inhibition rescues differentiation defect in E-protein knockout cells.** (A) qRT-PCR analysis of Nodal pathway gene expression in knockout and parental ESC line differentiation. Data are mean±s.d.,  $n=3$  biological replicates. \* $P<0.05$ , \*\* $P<0.01$  and \*\*\* $P<0.001$ , two-tailed paired Student's  $t$ -test. (B) Immunostaining of parental and knockout cells at day 5 of differentiation ±10 μM SB43. Scale bars: 50 μm. (C) Quantification of immunostaining of cells at day 5 of differentiation ±10 μM SB43 performed by quantitative image analysis. Data are the mean values for three biological replicates. A minimum of 8000 nuclei were scored per experiment. Data for all replicates are shown in Table S3.

neurodevelopmental deficiencies, and attempts to generate E-protein compound knockouts to overcome the expected problem of functional compensation between family members have thus far been unsuccessful as even mice with only single E-proteins knocked out do not survive beyond 2 weeks after birth

(Ravanpay and Olson, 2008). In *Drosophila*, however, loss of the only Class I bHLH factor gene, *daughterless*, has been found to result in neural differentiation defects (Caudy et al., 1988).

In this study we explored a role for E2A in neural fate commitment. Generation of an endogenously tagged E2A-V5

ESC line enabled us to define the expression of E2A from pluripotency to the onset of differentiation on the single cell level. Using a doxycycline-inducible system, we report that overexpression of E2A is sufficient to promote commitment to the neural lineage in ESCs, even under non-permissive culture conditions. Specifically, we found that forced E2A homodimers, rather than E2A-bHLH heterodimers, appear to drive this process, as the introduction of a single amino acid change (C570A) into E2A to disrupt homodimer stability, while still maintaining heterodimerisation capacity (Benezra, 1994), did ablate the ability of the forced homodimer to promote *Sox1* expression. These findings suggest that Id proteins inhibit neural differentiation, at least in part, by suppressing homodimerisation of E proteins: this would to our knowledge be the first example of a physiological role for E2A homodimers outside of the context of B-cell specification (Shen and Kadesch, 1995). It would be interesting to test this further by investigating whether E2A forced homodimers render neural differentiation resistant to the anti-neural effects of Id proteins, and whether introducing a C570A mutation into endogenous E2A recapitulates the neural defects seen in E2A null cells.

Interestingly, we also found that overexpression of E2A was able to drive upregulation of its own negative regulator, *Id1*, consistent with previous descriptions of a Class I/Class V HLH feedback loop in *Drosophila*, in which overexpression of the E-protein *daughterless* causes transcriptional upregulation of the Id protein Extra macrochaetae (Bhattacharya and Baker, 2011). We recognise that our forced homodimerisation of E2A is likely to disrupt this negative-feedback loop, and that this effect might help amplify the pro-neural activity of E2A homodimers.

The findings from the gain-of-function assays were further supported by data from the loss-of-function studies. Generation of novel *E2A*<sup>-/-</sup> and *E2A*<sup>-/-</sup>*HEB*<sup>-/-</sup> ESC lines highlighted that although knockout cells were able to self-renew normally, they were not able to differentiate efficiently. Additionally, although knockout cells were able to downregulate *Nanog*, they maintained relatively high expression of the epiblast markers *Fgf3* and *Oct4*. This suggests that although knockout cells are able to navigate the exit from pluripotency efficiently, they fail to complete the proposed second stage of neural lineage commitment involving the transition from epiblast-like cells to neuroectoderm-like cells (Zhang et al., 2010). The observation that *E2A*<sup>-/-</sup>*HEB*<sup>-/-</sup> cells displayed a more pronounced neural differentiation defect suggests that HEB is, indeed, able to at least partially compensate for the loss of E2A in this context. This is in line with previous studies that have demonstrated that HEB is able to functionally replace E2A during B-cell commitment (Zhuang et al., 1998), highlighting the extent of redundancy between E-protein family members. That *E2A*<sup>-/-</sup>*HEB*<sup>-/-</sup> cells also appear to preferentially upregulate markers of mesendodermal lineages, including *T* and *Eomes*, is a potentially interesting topic of investigation for future studies, especially considering that HEB (and E2A) have previously been implicated in promoting mesendodermal lineage specification, both in *Xenopus* and in human ESCs (Yoon et al., 2011; Li et al., 2017).

In this study, we identified *Foxd4* as a key gene strongly upregulated in response to forced E2A homodimers, positioning it upstream of *Sox1* expression during neural commitment. These data are in line with recent observations also made in ESCs (Sherman et al., 2017), and with the reported expression of *Foxd4* in the mouse neuroectoderm at E7.5 (Kaestner et al., 1995). The *Xenopus* homologue of *Foxd4*, *Foxd41l* (previously *Foxd5*), has a well-established role as part of a broader network of neural fate stabilising

factors (Yan and Moody, 2009), and it has been shown to expand the population of progenitor cells in the immature neuroectoderm, while repressing the transcription of genes associated with neural differentiation (Sullivan et al., 2001). A potential role for *Foxd4* in maintaining neuroectodermal cells in a proliferative non-differentiating state, therefore, might also explain the large proliferation of Oct4<sup>+</sup> cells we observed when *Foxd4* was ectopically expressed in *E2A*<sup>-/-</sup> cells.

We also reported that components of the Nodal signalling pathway were downregulated upon the overexpression of E2A homodimers, and conversely upregulated during the differentiation of E-protein knockout cells. Interestingly, E2A has also been shown to repress the transcription of *lefty* in *Xenopus*, with knockout of E2A causing upregulation of *lefty* and a subsequent failure in mesendodermal fate specification (Wills and Baker, 2015). We propose that E2A could be playing a similar role to repress Nodal signalling in mouse pluripotent cells, but during neural fate commitment – a process in which the inhibition of Nodal signalling is already known to be important in both mouse and human pluripotent cells (Watanabe et al., 2005; Vallier et al., 2009). Given that the inhibition of Nodal also promotes neural differentiation of human pluripotent cells (Vallier et al., 2004; Chambers et al., 2009) and that E2A regulates *lefty* in human cells (Yoon et al., 2011), it seems likely that E2A might also drive neural differentiation of human pluripotent cells, although this remains to be tested. It is likely that E2A has multiple downstream effectors in mammalian cells, as it does in *Xenopus* (Wills and Baker, 2015), but our rescue experiments suggest that the inhibition of Nodal and the activation of *Foxd4* are key effectors that explain the ability of E2A to drive neural differentiation. In summary, we propose that E2A plays an instructive, rather than passive, role in promoting neural fate commitment in pluripotent cells by promoting the transcription of neural lineage genes while simultaneously suppressing the Nodal signalling pathway.

## MATERIALS AND METHODS

### Mouse ESC culture and neural differentiation

E14tg2α mouse ESCs were used as wild-type cells. All cell lines used in this study were screened for mycoplasma contamination. ESCs were maintained in Glasgow minimum essential medium supplemented with β-mercaptoethanol, non-essential amino acids, glutamine, pyruvate, 10% fetal calf serum (FCS) and 100 units/ml LIF on gelatinised tissue culture flasks (Smith, 1991). Monolayer neural differentiation was performed as described by Pollard et al. (2006). Briefly, ESCs were washed in Dulbecco's modified eagle medium (DMEM)/F12 to remove all traces of serum and plated at 1×10<sup>4</sup> cells/cm<sup>2</sup> in N2B27 medium on gelatinised tissue culture plates. N2B27 consists of a 1:1 ratio of DMEM/F12 and Neurobasal media supplemented with 0.5% N2, 0.5% B27, L-glutamine and β-mercaptoethanol. For Nodal signalling inhibition experiments, cells were seeded in N2B27 supplemented with 10 μM SB431542.

### CRISPR/Cas9 epitope tagging of ESC lines

Guide RNA was manually designed to introduce the V5 tag into the 3' UTR of E2A based on the annotation of the final coding exon and the 3' UTR sequence. A ~200 bp sequence around the stop codon was used as an input for guide RNA design using either crispr.mit.edu or crispor.tefor.net web-based design tools. High-scoring guide RNAs were chosen based on minimal predicted off-target cleavage events, and having a cut site within the 3' UTR, preferably within 10 bp of the stop codon. For ssODN design, the V5 tag sequence was flanked by ~75 bp homology arms and a PAM-blocking mutation (NGG>NGC) was introduced into the 3' UTR sequence to prevent re-cutting of donor DNA by Cas9. Ribonucleoprotein (RNP) complexes were assembled as described by Dewari et al. (2018). Briefly, crRNA and tracrRNA oligos were mixed in equimolar concentration, heated

at 95°C for 5 min and allowed to cool to room temperature to anneal. Recombinant Cas9 protein (10 µg) was added to the annealed cr/tracrRNAs to form the RNP complex, which was incubated at room temperature for 20 min and stored on ice until electroporation. ssODN (30 pmol) was added to the RNP complex immediately before electroporation. ESCs ( $5 \times 10^4$ ) were transfected using the 4D Amaxa nucleofection system (Lonza) using the optimised program CA-210. Following transfection, cells were transferred into a six-well plate and allowed to recover for 3 to 5 days in ESC media. Clonal cell lines were isolated from bulk populations by manual colony picking and were subsequently analysed for successful knock-in by PCR genotyping, immunocytochemistry and Sanger sequencing of a ~500 bp region spanning the target site.

Synthetic Alt-R crRNA, tracrRNA, ssODN Ultramer template DNA oligonucleotides and recombinant Cas9 protein were manufactured by Integrated DNA Technologies. Guide RNA and ssODN DNA sequences are provided in Table S1.

### Knockout ESC line generation

For generation of *E2A*<sup>-/-</sup> ESCs, guide RNAs were designed to target exon 3 of the *E2A* gene. Targeting was performed on the *E2A*-V5 cell line to facilitate knockout line validation by loss of V5 signal. For generation of *E2A*<sup>-/-</sup>*HEB*<sup>-/-</sup> double knockout ESCs, exon 9 of the *HEB* locus was targeted in *E2A*<sup>-/-</sup> knockout cells (clone 9) in order to disrupt both *HEB*can and *HEB*alt splice variants. Guides were designed, assembled as RNPs with recombinant Cas9 and transfected as detailed above. Knockouts were verified by Sanger sequencing of targeted allele and western blot analysis using mouse anti-V5 (1:1000; Thermo Scientific, R960-25) and mouse anti-*HEB* (1:200; Santa Cruz Biotechnology, sc-28364) antibodies. Experiments presented in the main text were performed using *E2A*<sup>-/-</sup> clone 9 and *E2A*<sup>-/-</sup>*HEB*<sup>-/-</sup> clone 12.

### Doxycycline-inducible *E2A* cell lines

The E47 monomer construct comprises an N-terminally FLAG-tagged mouse E47 sequence. To generate the E47 forced homodimer construct, the DNA sequence corresponding to the FLAG-tagged mouse E47 was tethered at its C-terminus to the N-terminus of a second mouse E47 sequence with a 13-amino acid flexible linker sequence of TGSTGSKTGSTGS by overlapping extension PCR (Malaguti et al., 2019). For generation of the C570A homodimerisation-deficient cell lines, mutations were introduced into both E47 sequences in the forced dimer construct using PCR site-directed mutagenesis and verified by Sanger sequencing. Inducible cell lines were generated using an inducible cassette exchange technique (Iacovino et al., 2013). Cells were stimulated with 1 µg/ml doxycycline in all experiments.

### Generation of *Foxd4* rescue lines

The full-length *Foxd4*cDNA sequence was cloned into a pPGK-IRES-puromycin resistance plasmid. *E2A*<sup>-/-</sup> (clone 9) cells were transfected with the resulting construct using the 4D Amaxa nucleofection device (Lonza) and clonal lines were isolated by puromycin selection (1 µg/ml). Following clonal line expansion, ESCs were characterised by qRT-PCR for *Foxd4* expression.

### Immunocytochemistry, western blot analysis and flow cytometry

For immunocytochemistry, samples were fixed with 4% paraformaldehyde for 10 min at room temperature and incubated with blocking buffer (PBS, 0.1% Triton X-100, 3% donkey serum) for 30 min at room temperature. Primary antibodies were diluted in blocking buffer and incubated with cells overnight at 4°C. After three washes with PBS, cells were incubated with secondary antibodies conjugated with Alexa fluorophores (Life Technologies) diluted 1:1000 in blocking buffer for 1 h at room temperature. For nuclear counterstaining, cells were incubated with 1 µg/ml DAPI (Sigma-Aldrich) for 5 min following immunostaining. Cells were washed a minimum of three times in PBS before imaging. For confocal imaging, cells were plated onto glass coverslips, stained and mounted in ProLong Gold Antifade Reagent (Life Technologies). Cells were either imaged on the Leica SP8 confocal microscope or the Olympus IX-81 widefield microscope. All images were analysed in Fiji. Primary antibodies used for immunocytochemistry

were: mouse anti-β-III-tubulin/Tuj1 (1:1000; BioLegend, MMS-435P), rat anti-Nanog (1:200; Thermo Fisher Scientific, 14-5761-80), goat anti-Oct4 (1:200; Santa Cruz, sc-8628), mouse anti-Sox1 (1:200; BD Pharmingen, 560749), rabbit anti-Sox2 (1:200; Abcam, ab97959) and mouse anti-V5 (1:200-500; Thermo Fisher Scientific, R960-25). Secondary antibodies used in immunostaining experiments were: donkey anti-mouse Alexa Fluor 488 (Thermo Scientific, A21202), donkey anti-mouse Alexa Fluor 568 (Thermo Scientific, A10037), donkey anti-rabbit Alexa Fluor 488 (Thermo Scientific, A21206), donkey anti-rabbit 568 (Thermo Scientific, A10042), donkey anti-rat Alexa Fluor 488 (Thermo Scientific, A21208), donkey anti-rat Alexa Fluor 594 (Thermo Scientific, A21209) and donkey anti-goat Alexa Fluor 647 (Thermo Scientific, A21447).

For western blot analysis, cells were lysed in radioimmunoprecipitation assay buffer+1× PMSF (Alpha Diagnostics). Protein lysate (20 µg) was run on 4%-12% NuPAGE Bis-Tris Gel (Novex) and transferred onto Amersham Hybond ECL Nitrocellulose Membrane (GE Healthcare). Membranes were blocked in 5% Amersham ECL Prime Blocking Agent (GE Healthcare)+0.1% Tween 20 (Sigma-Aldrich) in PBS for a minimum of 1 h at room temperature. Membranes were incubated with primary antibody overnight at 4°C, washed in PBS+0.1% Tween 20, incubated in horseradish peroxidase (HRP)-conjugated secondary antibody for 1 h at room temperature and washed in PBS+0.1% Tween 20. Membranes were subsequently incubated in Amersham ECL Western Blotting Detection Reagent (GE Healthcare) or Amersham ECL Prime Western Blotting Detection Reagent (GE Healthcare). Membranes were either used to expose Amersham Hyperfilm ECL (GE Healthcare) and developed using a Konica SRX-101A Medical Film Processor, or imaged using the Bio-Rad ChemiDoc system. Primary antibodies used in the western blot analysis were mouse anti-β-III-tubulin/Tuj1 (1:1000; Sigma-Aldrich, T5293), mouse anti-E2A (1:200; Santa Cruz Biotechnology, sc-416), mouse anti-FLAG (1:5000; Sigma-Aldrich, F9291), mouse anti-*HEB* (1:200; Santa Cruz Biotechnology, sc-28364), mouse anti-Sox1 (1:1000; BD Pharmingen, 560749) and V5 (1:500-1:1000; Thermo Fisher Scientific, R960-25). Secondary antibodies used were Amersham ECL Mouse IgG, HRP linked (GE Healthcare, NA931V) and Amersham ECL Rabbit IgG, HRP linked (GE Healthcare, NA934).

For flow cytometry analysis, cells were dissociated into a single cell suspension in PBS+10% FCS. Cells were stained with rat anti-E-cadherin (CD324), eFluor660-conjugated DECMA-1 (1:300; eBioscience, 50-3249-82) and 100 ng/ml DAPI to stain dead cells. Flow cytometry was performed on the BD Accuri C6 and analysis was performed using FlowJo software.

### Quantification of immunostaining

Immunofluorescence was quantified by nuclear segmentation based on DAPI signal and manual editing of segmentation results using Nessys software, as detailed by Blin et al. (2019).

### qRT-PCR

Total RNA was isolated from cells using the Absolutely RNA Miniprep Kit (Agilent). cDNA was generated using Moloney murine leukaemia virus reverse transcriptase and random primers. Primers and UPL probes (Roche) used are detailed in Table S2. All gene expression values were normalised to the housekeeping gene *Sdha*.

### RNA-seq

RNA was isolated from cells using the Absolutely RNA Miniprep Kit (Agilent) and RNA quality was verified using an Agilent 2100 Bioanalyzer. Subsequent cDNA library preparation, sequencing and bioinformatics analysis, including differential gene expression analyses, were performed by the Edinburgh Genomics facility. Library preparation was performed using the TruSeq Stranded mRNA Library Prep Kit (Illumina) and libraries were sequenced on the Illumina HiSeq4000. Reads were mapped to the mouse genome GRCm38 from Ensembl and were aligned to the reference genome using STAR (version 2.5.2b). Differential gene analysis was carried out with edgeR (version 3.16.5). Differentially expressed genes were assigned based on a minimum fold-change of two and a maximum FDR of 0.05. Gene ontology analysis was performed on genes upregulated at 24 h using the STRING database (Szklarczyk et al., 2017).



## Statistical analyses

qRT-PCR data are represented as mean±s.d. for a minimum of three experimental replicates. Statistical significance was calculated using Student's *t*-tests for pairwise comparisons, or ANOVA with correction for multiple comparisons for two or more samples. Statistically significant differences are shown as follows: \**P*<0.05, \*\**P*<0.01 and \*\*\**P*<0.001.

## Acknowledgements

We thank Steve Pollard and Pooran Dewari for their help with designing the CRISPR/Cas9 targeting experiments; Thomas Maynard for providing the Foxd4 cDNA; and Edinburgh Genomics for performing library preparation, RNA-sequencing and bioinformatics analysis.

## Competing interests

The authors declare no competing or financial interests.

## Author contributions

Conceptualization: C.R., J.O.M., S.L.; Methodology: C.R., M.M.; Formal analysis: C.R., M.M., S.L.; Investigation: C.R., M.M.; Resources: M.M., S.L.; Data curation: M.M.; Writing - original draft: C.R.; Writing - review & editing: C.R., M.M., J.O.M., S.L.; Visualization: C.R.; Supervision: M.M., J.O.M., S.L.; Project administration: S.L.; Funding acquisition: S.L.

## Funding

This study was supported by the Wellcome Trust (WT103789AIA to S.L.), and by PhD studentship funding to C.R. from the Medical Research Council (MR/K501293/1) and the Wellcome Trust (108906/Z/15/Z). Open access funding provided by the University of Edinburgh. Deposited in PMC for immediate release.

## Data availability

RNA-seq data have been deposited in GEO under accession number GSE152118.

## Supplementary information

Supplementary information available online at

<https://dev.biologists.org/lookup/doi/10.1242/dev.184093.supplemental>

## Peer review history

The peer review history is available online at

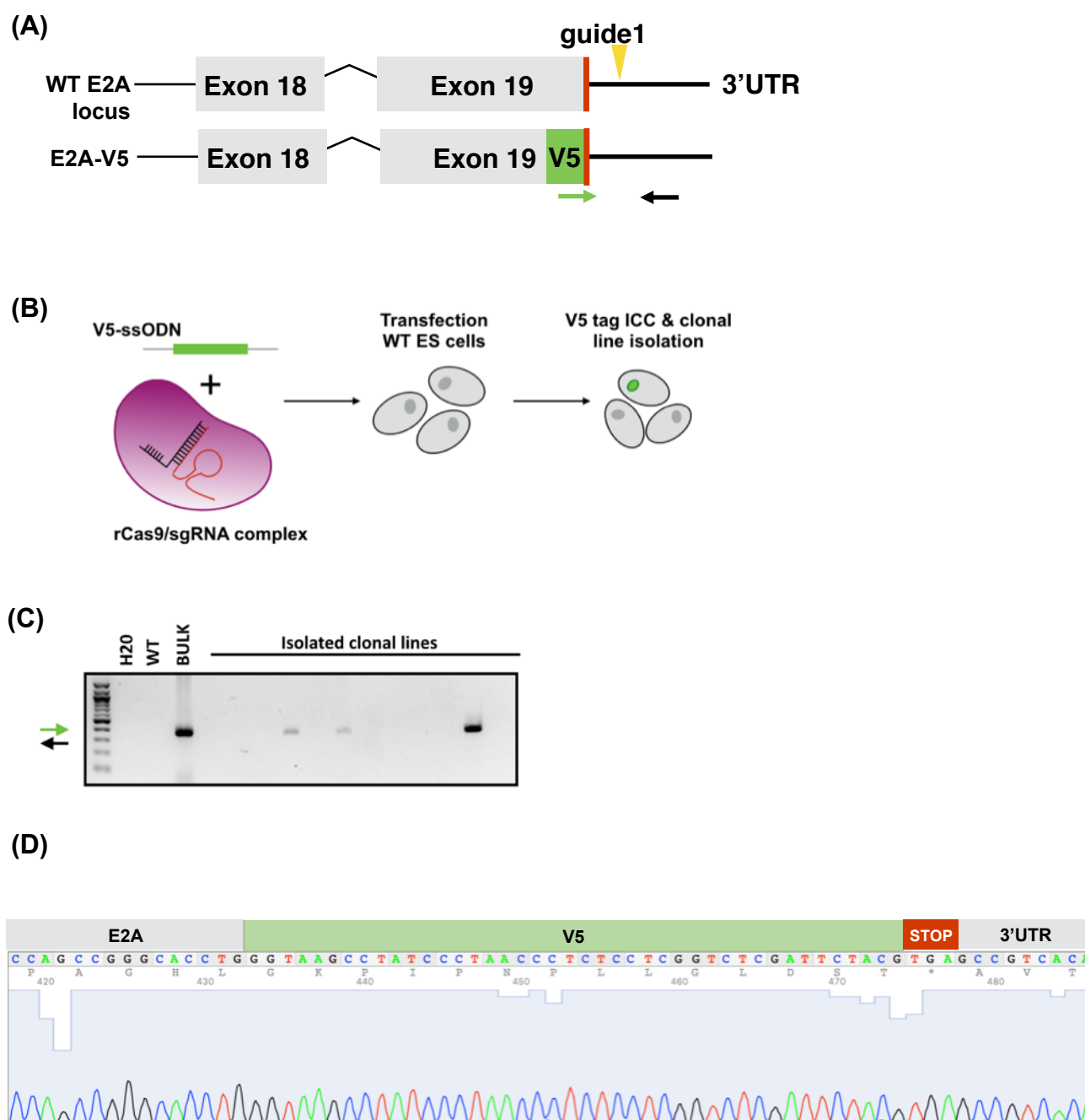
<https://dev.biologists.org/lookup/doi/10.1242/dev.184093.reviewer-comments.pdf>

## References

- Aiba, K., Nedorezov, T., Piao, Y., Nishiyama, A., Matoba, R., Sharova, L. V., Sharov, A. A., Yamanaka, S., Niwa, H. and Ko, M. S. H. (2009). Defining developmental potency and cell lineage trajectories by expression profiling of differentiating mouse embryonic stem cells. *DNA Res.* **16**, 73–80. doi:10.1093/dnares/dsn035
- Aloia, L., Gutierrez, A., Caballero, J. M. and Di Croce, L. (2015). Direct interaction between Id1 and Zrf1 controls neural differentiation of embryonic stem cells. *EMBO Rep.* **16**, 63–70. doi:10.15252/embr.201439560
- Bain, G., Maandag, E. C., Izon, D. J., Amsen, D., Kruisbeek, A. M., Weintraub, B. C., Krop, I., Schlissel, M. S., Feeney, A. J. and van Roon, M. (1994). E2A proteins are required for proper B cell development and initiation of immunoglobulin gene rearrangements. *Cell* **79**, 885–892. doi:10.1016/0092-8674(94)90077-9
- Benezra, R. (1994). An intermolecular disulfide bond stabilizes E2A homodimers and is required for DNA binding at physiological temperatures. *Cell* **79**, 1057–1067. doi:10.1016/0092-8674(94)90036-1
- Bhattacharya, A. and Baker, N. E. (2011). A network of broadly expressed HLH genes regulates tissue-specific cell fates. *Cell* **147**, 881–892. doi:10.1016/j.cell.2011.08.055
- Blin, G., Sadurska, D., Portero Migueles, R., Chen, N., Watson, J. A. and Lowell, S. (2019). Nssys: a new set of tools for the automated detection of nuclei within intact tissues and dense 3D cultures. *PLoS Biol.* **17**, e3000388. doi:10.1371/journal.pbio.3000388
- Camus, A., Perea-Gomez, A., Moreau, A. and Collignon, J. (2006). Absence of Nodal signaling promotes precocious neural differentiation in the mouse embryo. *Dev. Biol.* **295**, 743–755. doi:10.1016/j.ydbio.2006.03.047
- Caudy, M., Vassin, H., Brand, M., Tuma, R., Jan, L. Y. and Jan, Y. N. (1988). daughterless, a drosophila gene essential for both neurogenesis and sex determination, has sequence similarities to myc and the achaete-scute complex. 1–7. *Cell* **55**, 1061–1067. doi:10.1016/0092-8674(88)90250-4
- Chambers, S. M., Fasano, C. A., Papapetrou, E. P., Tomishima, M., Sadelain, M. and Studer, L. (2009). Highly efficient neural conversion of human ES and iPS cells by dual inhibition of SMAD signaling. *Nat. Biotechnol.* **27**, 275–280. doi:10.1038/nbt.1529
- Davies, O. R., Lin, C.-Y., Radzisheuskaya, A., Zhou, X., Taube, J., Blin, G., Waterhouse, A., Smith, A. J. H. and Lowell, S. (2013). Tcf15 primes pluripotent cells for differentiation. *Cell Rep.* **3**, 472–484. doi:10.1016/j.celrep.2013.01.017
- Dewari, P. S., Southgate, B., McCarten, K., Monogarov, G., O'Duibhir, E., Quinn, N., Tyrer, A., Leitner, M.-C., Plumb, C., Kalantzaki, M. et al. (2018). An efficient and scalable pipeline for epitope tagging in mammalian stem cells using Cas9 ribonucleoprotein. *Elife* **7**, e35069. doi:10.7554/eLife.35069
- Di-Gregorio, A., Sancho, M., Stuckey, D. W., Crompton, L. A., Godwin, J., Mishina, Y. and Rodriguez, T. A. (2007). BMP signalling inhibits premature neural differentiation in the mouse embryo. *Development* **134**, 3359–3369. doi:10.1242/dev.005967
- Haeghele, L., Ingold, B., Naumann, H., Tabatabai, G., Ledermann, B. and Brandner, S. (2003). Wnt signalling inhibits neural differentiation of embryonic stem cells by controlling bone morphogenetic protein expression. *Mol. Cell. Neurosci.* **24**, 696–708. doi:10.1016/S1044-7431(03)00232-X
- Harland, R. (2000). Neural induction. *Curr. Opin. Genet. Dev.* **10**, 357–362. doi:10.1016/S0959-437X(00)00096-4
- Iacovino, M., Roth, M. E. and Kyba, M. (2013). Rapid genetic modification of mouse embryonic stem cells by inducible cassette exchange recombination. In *Embryonic Stem Cell Protocols* (ed. K. Turksen), pp. 339–351. Totowa, NJ: Humana Press.
- Imayoshi, I. and Kageyama, R. (2014). bHLH factors in self-renewal, multipotency, and fate choice of neural progenitor cells. *Neuron* **82**, 9–23. doi:10.1016/j.neuron.2014.03.018
- Inman, G. J., Nicolás, F. J., Callahan, J. F., Harling, J. D., Gaster, L. M., Reith, A. D., Laping, N. J. and Hill, C. S. (2002). SB-431542 is a potent and specific inhibitor of transforming growth factor- $\beta$  superfamily type I activin receptor-like kinase (ALK) receptors ALK4, ALK5, and ALK7. *Mol. Pharmacol.* **62**, 65–74. doi:10.1124/mol.62.1.65
- Iwafuchi-Doi, M., Matsuda, K., Murakami, K., Niwa, H., Tesar, P. J., Aruga, J., Matsuo, I. and Kondoh, H. (2012). Transcriptional regulatory networks in epiblast cells and during anterior neural plate development as modeled in epiblast stem cells. *Development* **139**, 4675–4675. doi:10.1242/dev.091181
- Kaestner, K. H., Monaghan, A. P., Kern, H., Ang, S.-L., Weitz, S., Lichter, P. and Schütz, G. (1995). The Mouse fkh-2 Gene implications for notochord, foregut, and midbrain regionalization. *J. Biol. Chem.* **270**, 30029–30035. doi:10.1074/jbc.270.50.30029
- Kamiya, D., Banno, S., Sasai, N., Ohgushi, M., Inomata, H., Watanabe, K., Kawada, M., Yakura, R., Kiyonari, H., Nakao, K. et al. (2011). Intrinsic transition of embryonic stem-cell differentiation into neural progenitors. *Nature* **470**, 503–509. doi:10.1038/nature09726
- Koyano-Nakagawa, N., Kim, J., Anderson, D. and Kintner, C. (2000). Hes6 acts in a positive feedback loop with the neurogenins to promote neuronal differentiation. *Development* **127**, 4203–4216.
- Li, Y., Brauer, P. M., Singh, J., Xhiku, S., Yoganathan, K., Zúñiga-Pflücker, J. C. and Anderson, M. K. (2017). Targeted disruption of TCF12 reveals HEB as essential in human mesodermal specification and hematopoiesis. *Stem Cell Reports* **9**, 779–795. doi:10.1016/j.stemcr.2017.07.011
- Lin, C.-Y., Tatar, T., Blin, G., Malaguti, M., Migueles, R. P., Shao, H., Chen, N., Chambers, I. and Lowell, S. (2019). Competence to epithelialise coincides with competence to differentiate in pluripotent cells. *bioRxiv* **17**, e3000388-e26. doi:10.1101/809467
- Malaguti, M., Nistor, P. A., Blin, G., Pegg, A., Zhou, X. and Lowell, S. (2013). Bone morphogenetic protein signalling suppresses differentiation of pluripotent cells by maintaining expression of E-Cadherin. *Elife* **2**, e01197. doi:10.7554/eLife.01197
- Malaguti, M., Migueles, R. P., Blin, G., Lin, C.-Y. and Lowell, S. (2019). Id1 stabilizes epiblast identity by sensing delays in nodal activation and adjusting the timing of differentiation. *Dev. Cell* **50**, 462–477. e465. doi:10.1016/j.devcel.2019.05.032
- Murai, K., Vernon, A. E., Philpott, A. and Jones, P. (2007). Hes6 is required for MyoD induction during gastrulation. *Dev. Biol.* **312**, 61–76. doi:10.1016/j.ydbio.2007.09.011
- Murre, C., McCaw, P. S. and Baltimore, D. (1989). A new DNA binding and dimerization motif in immunoglobulin enhancer binding, daughterless, MyoD, and myc proteins. *Cell* **56**, 777–783. doi:10.1016/0092-8674(89)90682-X
- Neuhof, L. A. and Wold, B. (1993). HLH forced dimers: tethering MyoD to E47 generates a dominant positive myogenic factor insulated from negative regulation by Id. *Cell* **74**, 1033–1042. doi:10.1016/0092-8674(93)90725-6
- Norton, J. D. (2000). ID helix-loop-helix proteins in cell growth, differentiation and tumorigenesis. *J. Cell Sci.* **113**, 3897–3905.
- Peng, G., Suo, S., Chen, J., Chen, W., Liu, C., Yu, F., Wang, R., Chen, S., Sun, N., Cui, G. et al. (2016). Spatial transcriptome for the molecular annotation of lineage fates and cell identity in mid-gastrula mouse embryo. *Dev. Cell* **36**, 681–697. doi:10.1016/j.devcel.2016.02.020
- Pérez-Moreno, M. A., Locascio, A., Rodrigo, I., Dhondt, G., Portillo, F., Nieto, M. A. and Cano, A. (2001). A new role for E12/E47 in the repression of e-cadherin expression and epithelial-mesenchymal transitions. *J. Biol. Chem.* **276**, 27424–27431. doi:10.1074/jbc.M100827200

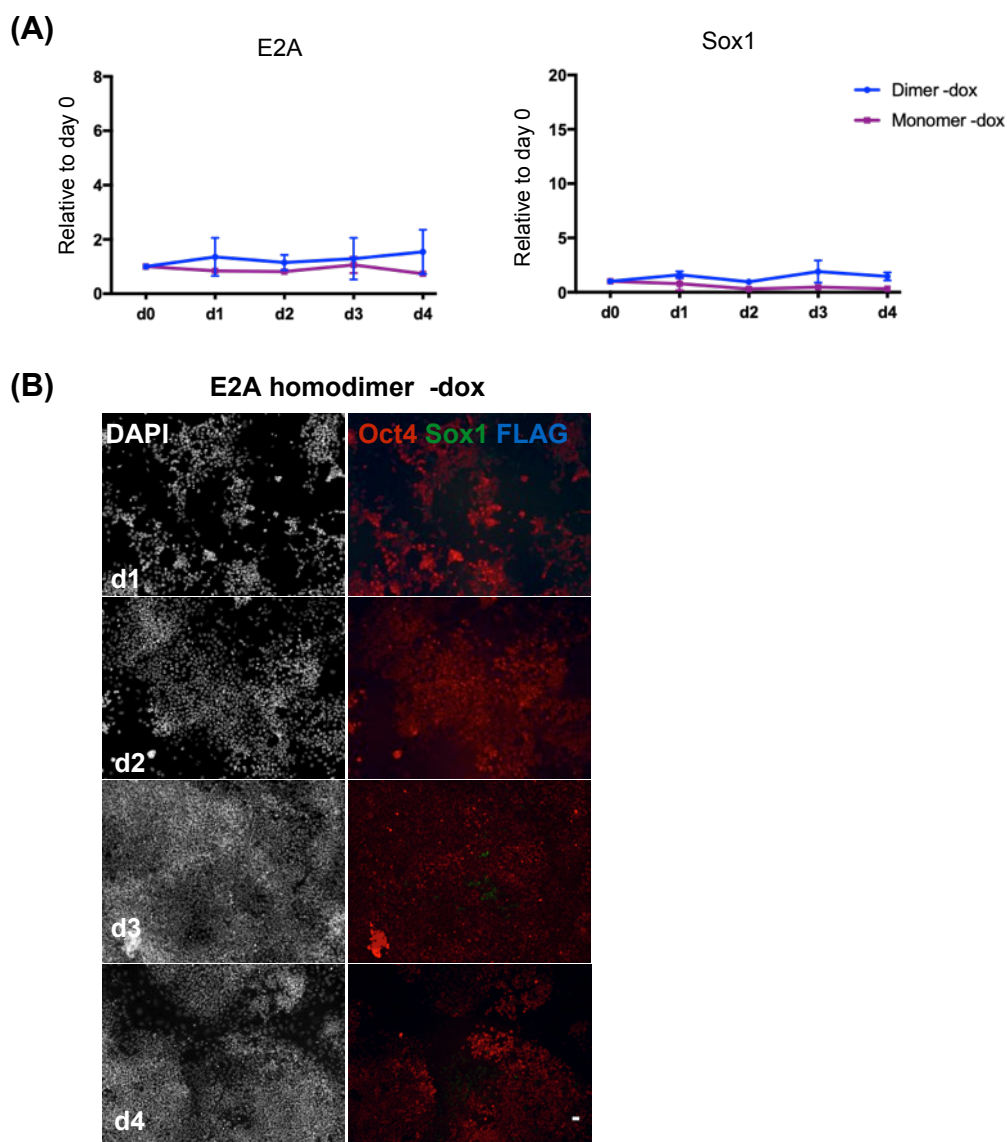


- Pevny, L. H., Sockanathan, S., Placzek, M. and Lovell-Badge, R. (1998). A role for SOX1 in neural determination. *Development* **125**, 1967-1978.
- Pollard, S. M., Benchoua, A. and Lowell, S. (2006). Neural stem cells, neurons, and glia. *Methods Enzymol.* **418**, 151-169. doi:10.1016/S0076-6879(06)18010-6
- Ravanpay, A. C. and Olson, J. M. (2008). E protein dosage influences brain development more than family member identity. *J. Neurosci. Res.* **86**, 1472-1481. doi:10.1002/jnr.21615
- Sankar, S., Yellajoshyula, D., Zhang, B., Teets, B., Rockweiler, N. and Kroll, K. L. (2016). Gene regulatory networks in neural cell fate acquisition from genome-wide chromatin association of Geminin and Zic1. *Sci. Rep.* **6**, 37412. doi:10.1038/srep37412
- Shen, C. P. and Kadesch, T. (1995). B-cell-specific DNA binding by an E47 homodimer. *Mol. Cell. Biol.* **15**, 4518-4524. doi:10.1128/MCB.15.8.4518
- Sherman, J. H., Karpinski, B. A., Fralish, M. S., Cappuzzo, J. M., Dhindsa, D. S., Thal, A. G., Moody, S. A., LaMantia, A. S. and Maynard, T. M. (2017). Foxd4 is essential for establishing neural cell fate and for neuronal differentiation. *Genesis* **55**, e23031. doi:10.1002/dvg.23031
- Smith, A. G. (1991). Culture and differentiation of embryonic stem cells. *Journal of Tissue Culture Methods* **13**, 89-94. doi:10.1007/BF01666137
- Smith, J. R., Vallier, L., Lup, G., Alexander, M., Harris, W. A. and Pedersen, R. A. (2008). Inhibition of Activin/Nodal signaling promotes specification of human embryonic stem cells into neuroectoderm. *Dev. Biol.* **313**, 107-117. doi:10.1016/j.ydbio.2007.10.003
- Sullivan, S. A., Moody, S. A. and Akers, L. (2001). foxD5a, a Xenopus winged helix gene, maintains an immature neural ectoderm via transcriptional repression that is dependent on the C-terminal domain. *Dev. Biol.* **232**, 439-457. doi:10.1006/dbio.2001.0191
- Szklarczyk, D., Morris, J. H., Cook, H., Kuhn, M., Wyder, S., Simonovic, M., Santos, A., Doncheva, N. T., Roth, A., Bork, P. et al. (2017). The STRING database in 2017: quality-controlled protein-protein association networks, made broadly accessible. *Nucleic Acids Res.* **45**, D362-D368. doi:10.1093/nar/gkw937
- Tam, P. P. L. and Zhou, S. X. (1996). The allocation of epiblast cells to ectodermal and germ-line lineages is influenced by the position of the cells in the gastrulating mouse embryo. *Dev. Biol.* **178**, 124-132. doi:10.1006/dbio.1996.0203
- Vallier, L., Reynolds, D. and Pedersen, R. A. (2004). Nodal inhibits differentiation of human embryonic stem cells along the neuroectodermal default pathway. *Dev. Biol.* **275**, 403-421. doi:10.1016/j.ydbio.2004.08.031
- Vallier, L., Touboul, T., Chng, Z., Brimpari, M., Hannan, N., Millan, E., Smithers, L. E., Trotter, M., Rugg-Gunn, P., Weber, A. et al. (2009). Early cell fate decisions of human embryonic stem cells and mouse epiblast stem cells are controlled by the same signalling pathways. *PLoS ONE* **4**, e6082. doi:10.1371/journal.pone.0006082
- Watanabe, K., Kamiya, D., Nishiyama, A., Katayama, T., Nozaki, S., Kawasaki, H., Watanabe, Y., Mizuseki, K. and Sasai, Y. (2005). Directed differentiation of telencephalic precursors from embryonic stem cells. *Nat. Neurosci.* **8**, 288-296. doi:10.1038/nn1402
- Wills, A. E. and Baker, J. C. (2015). E2a is necessary for Smad2/3-dependent transcription and the direct repression of lefty during gastrulation. *Dev. Cell* **32**, 345-357. doi:10.1016/j.devcel.2014.11.034
- Yan, B. and Moody, S. A. (2009). foxD5 plays a critical upstream role in regulating neural ectodermal fate and the onset of neural differentiation. *Dev. Biol.* **329**, 80-95. doi:10.1016/j.ydbio.2009.02.019
- Ying, Q.-L., Nichols, J., Chambers, I. and Smith, A. (2003a). BMP induction of Id proteins suppresses differentiation and sustains embryonic stem cell self-renewal in collaboration with STAT3. *Cell* **115**, 281-292. doi:10.1016/S0092-8674(03)00847-X
- Ying, Q.-L., Stavridis, M., Griffiths, D., Li, M. and Smith, A. (2003b). Conversion of embryonic stem cells into neuroectodermal precursors in adherent monoculture. *Nat. Biotechnol.* **21**, 183-186. doi:10.1038/nbt780
- Yoon, S.-J., Wills, A. E., Chuong, E., Gupta, R. and Baker, J. C. (2011). HEB and E2A function as SMAD/FOXH1 cofactors. *Genes Dev.* **25**, 1654-1661. doi:10.1101/gad.16800511
- Zhang, K., Li, L., Huang, C., Shen, C., Tan, F., Xia, C., Liu, P., Rossant, J. and Jing, N. (2010). Distinct functions of BMP4 during different stages of mouse ES cell neural commitment. *Development* **137**, 2095-2105. doi:10.1242/dev.049494
- Zhu, Q., Song, L., Peng, G., Sun, N., Chen, J., Zhang, T., Sheng, N., Tang, W., Qian, C., Qiao, Y. et al. (2014). The transcription factor Pou3f1 promotes neural fate commitment via activation of neural lineage genes and inhibition of external signaling pathways. *Elife* **3**, e02224. doi:10.7554/eLife.02224.020
- Zhuang, Y., Soriano, P. and Weintraub, H. (1994). The helix-loop-helix gene E2A is required for B cell formation. *Cell* **79**, 875-884. doi:10.1016/0092-8674(94)90076-0
- Zhuang, Y., Barndt, R. J., Pan, L., Kelley, R. and Dai, M. (1998). Functional replacement of the mouse E2A gene with a human HEB cDNA. *Mol. Cell. Biol.* **18**, 3340-3349. doi:10.1128/MCB.18.6.3340



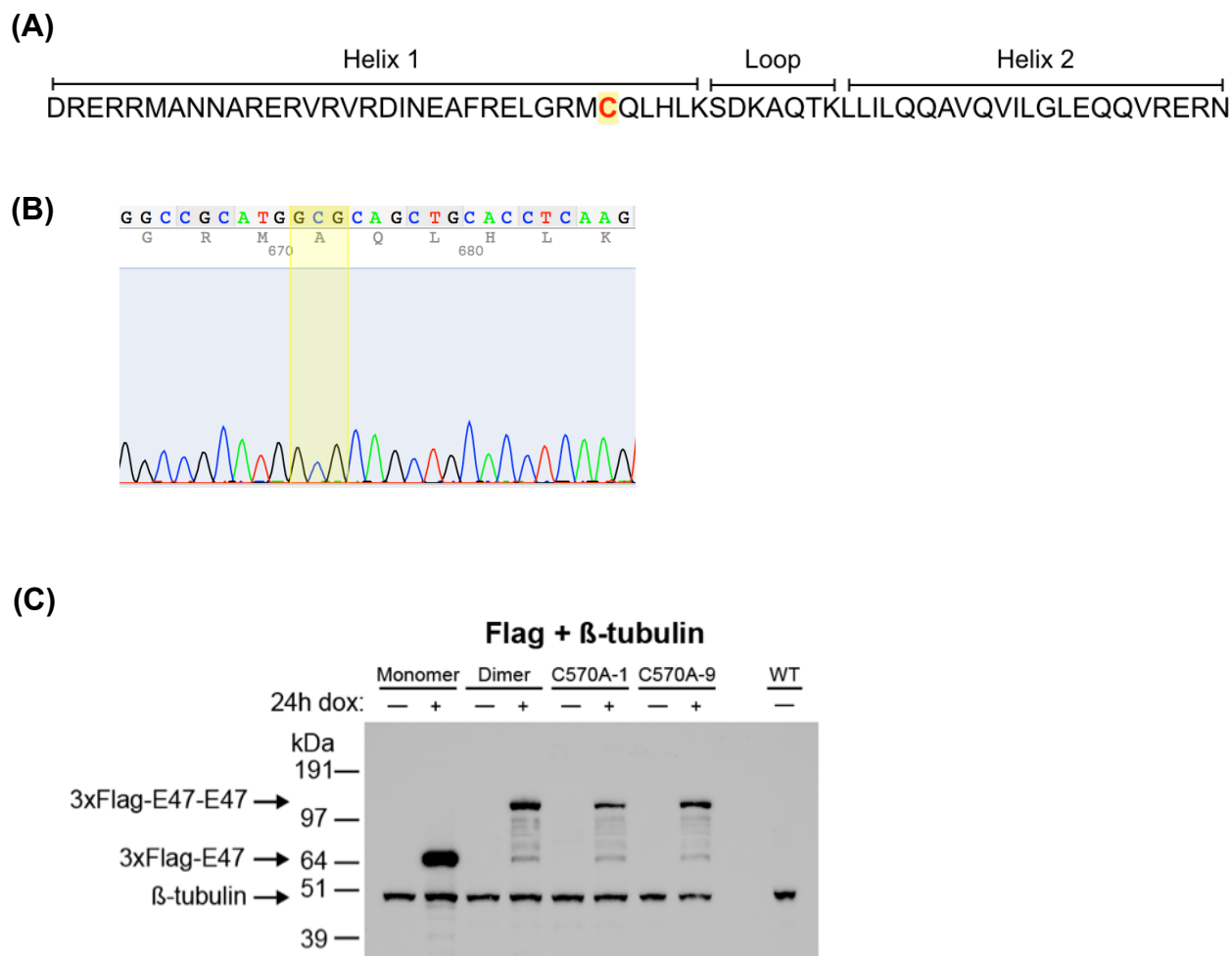
**Figure Q. Generation and validation of endogenously tagged E2A-V5 ES cell lines.**

- (A)** Schematic of exon targeting of E2A to generate tagged cell lines. V5 tag is shown in green, stop codon in red, and yellow arrow indicates location of the guide RNA (sgRNA) sequence.
- (B)** Method used to knock in the V5 tag by nucleofection of a complexed ribonucleoprotein (RNP) comprised of recombinant Cas9 protein (rCas9) and guide RNA into wild-type ES cells.
- (C)** PCR genotyping of derived clonal lines using primer pairs indicated by green and black arrows depicted in (A). Bands of correct size in clonal lines shown, compared to blank (H2O), parental WT line, and bulk transfected population.
- (D)** Sanger sequencing trace of correctly targeted E2A-V5 clonal lines, confirming correct in-frame insertion of the V5 tag.



**Figure Q2. Sox1 is not expressed when inducible cells are cultured in LIF/serum without dox induction.**

- (A)** qRT-PCR analysis of dox-inducible E2A monomer and forced homodimer cells cultured in LIF/serum without dox. Expression values are normalised to day 0. Error bars represent mean  $\pm$  SD of 2 biological replicates.
- (B)** Immunocytochemistry of inducible forced homodimer cells cultured in LIF/serum without dox. Scale bar: 30  $\mu$ m



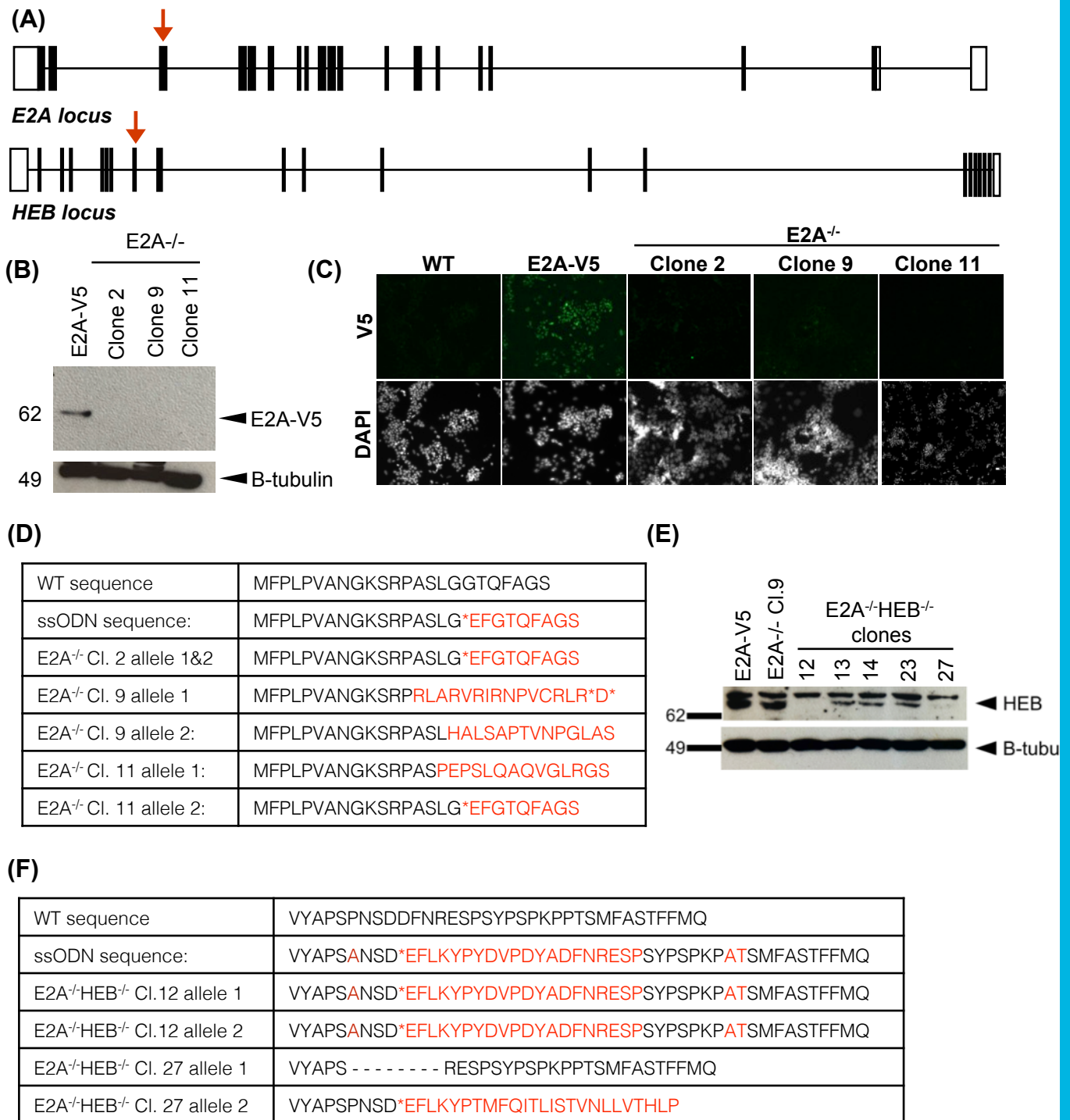
**Figure 3. Wild-type and mutant forced homodimer cells.**

**(A)** Sequence of the helix-loop-helix (HLH) domain of wild-type E2A with Cysteine-570 residue in helix 1 highlighted in yellow.

**(B)** Sanger sequencing trace of Cys>Ala (C570A) mutation introduced into both E2A monomer sequences in the forced homodimer construct to generate mutant forced homodimer inducible lines (C570A-1 and C570A-9).

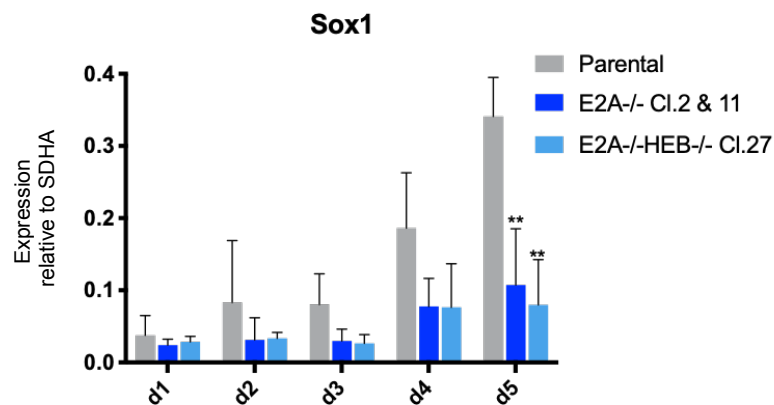
**(C)** Western blot analysis of wild type monomer, forced homodimer and mutant forced homodimer cell lines +/- 24h of doxycycline induction. Membrane was blotted using anti-FLAG to detect the transgene and B-tubulin was used as a loading control.



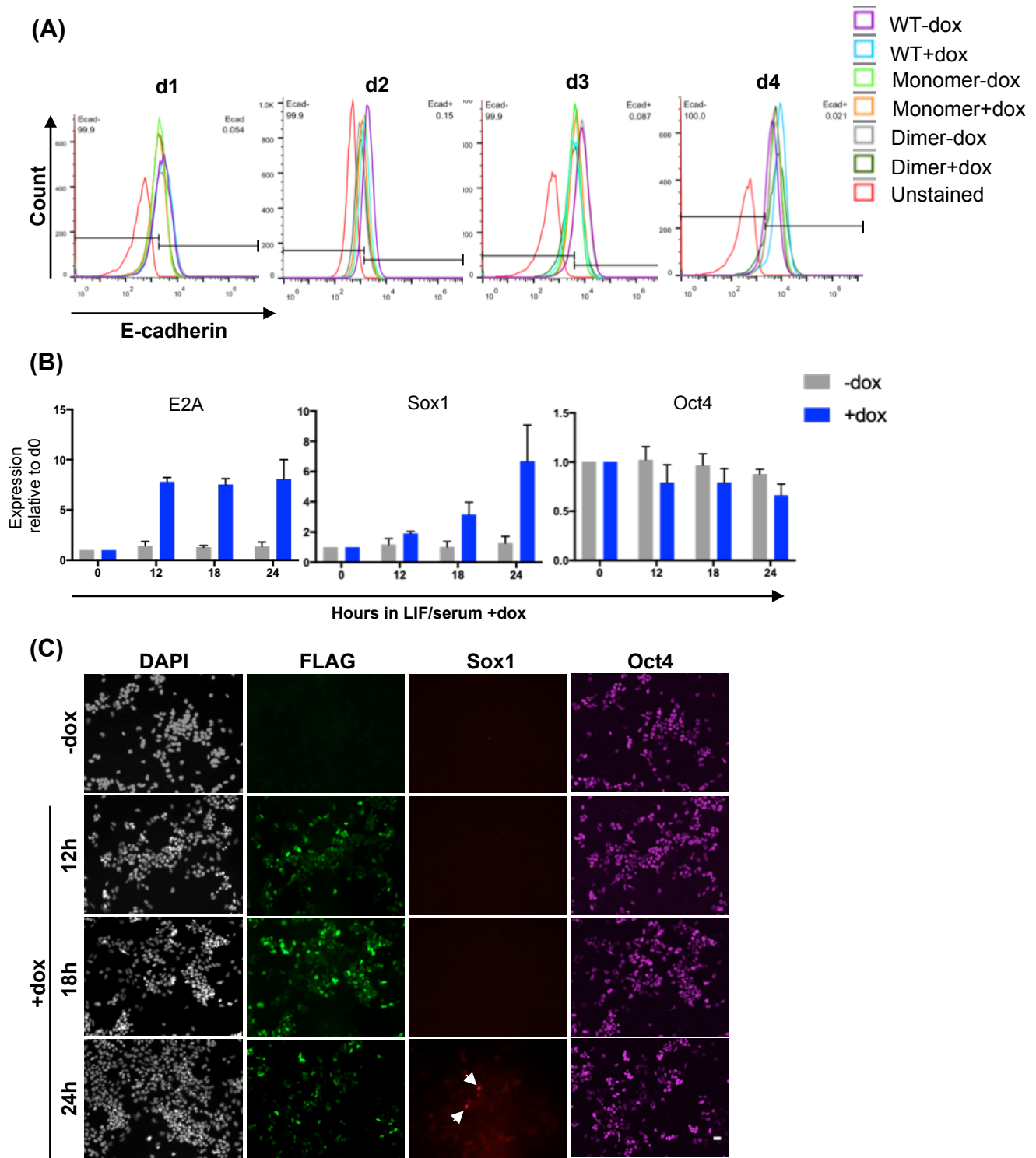


#### Figure C4. Validation of E-protein knockout lines

**(A)** Structure of *E2A* and *HEB* loci. Red arrow indicates Cas9/sgrNA target site. **(B)** Western blot of E2A-V5 Cl.74 (parental cell line) and three *E2A* knockout clones 2, 9 & 11 using anti-V5 antibody. E2A-V5=69kDa. B-tubulin was analysed as a loading control. **(C)** Immunostaining of WT, E2A-V5 (parental) and *E2A*<sup>-/-</sup> clonal lines with anti-V5 antibody. Scale bar: 30 μm **(D)** Sanger sequencing of *E2A*<sup>-/-</sup> clones 2, 9 and 11. WT exon 3 and ssODN template DNA amino acid sequences shown with mutations highlighted in red. *E2A*<sup>-/-</sup> clone 2 has biallelic knock-in of the stop codon-containing ssODN template sequence, clone 9 has two independently disrupted *E2A* alleles and clone 11 has heterozygous knock-in of the ssODN and a second non-homologous end joining (NHEJ)-disrupted allele. **(E)** Western blot of E2A-V5 Cl.74, *E2A*<sup>-/-</sup> clone 9 (parental) and *E2A*<sup>-/-</sup>HEB<sup>-/-</sup> clonal lines using anti-HEB antibody. HEB=85kDa. B-tubulin was analysed as a loading control. **(F)** Sanger sequencing of *E2A*<sup>-/-</sup>HEB<sup>-/-</sup> clones 12 and 27. Targeted WT exon 9 and ssODN template DNA amino acid sequences shown above, with mutations highlighted in red. *E2A*<sup>-/-</sup>HEB<sup>-/-</sup> clone 12 has biallelic knock-in of the ssODN sequence, and clone 27 has two independently disrupted *HEB* alleles; one allele contains an 8 amino acid deletion, the other allele contains a premature stop codon.

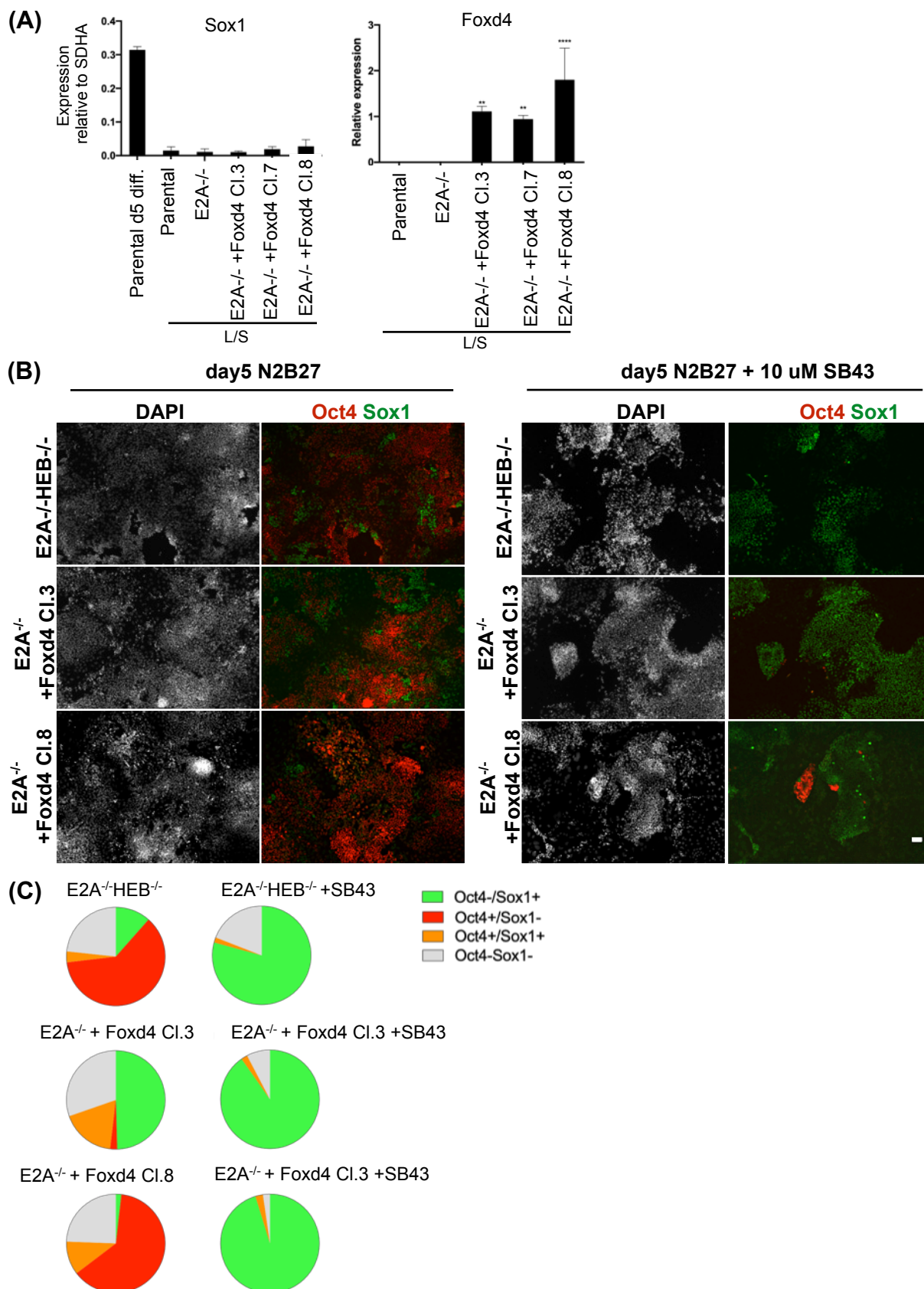


**Figure G5. Transcriptional analysis of additional E-protein knockout clones.** qRT-PCR analysis of two additional E2A<sup>-/-</sup> clones (clones 2 and 11) and one additional E2A<sup>-/-</sup>HEB<sup>-/-</sup> clone (clone 27) differentiated in N2B27. Relative expression values shown are the mean of three independent experiments for each clonal line.



**Figure C6. Identifying a time frame for detection of early E2A homodimer response genes.**

(A) Flow cytometric analysis of the effect of overexpression of E2A monomers and homodimers on E-cadherin (E-cad) protein expression. WT ES cells are included as a control and were used in combination with unstained WT cells to define gates. Experiments were repeated twice and the data from replicates follow the same trend as displayed in the figure. (B) qRT-PCR analysis of cells following 12, 18 and 24 hours of E2A homodimer induction in LIF/serum. No dox controls are also shown. Error bars represent standard deviation of three biological experiments, which were used for subsequent RNA-sequencing analysis. (C) Immunostaining of cells during the 24h timecourse to assess transgene activation using an anti-FLAG antibody, and co-staining with anti-Sox1 and anti-Oct4. White arrows highlight small number of cells that express Sox1 protein 24 hours post-induction. Scale bar: 50  $\mu$ m



**Figure G7. Analysis of additional clonal lines ectopically overexpressing Foxd4 in an E2A<sup>-/-</sup> background.**

(A) qRT-PCR analysis of knockout and Foxd4 rescue cell lines in LIF/serum. Relative expression values shown are the mean  $\pm$  SD of three biological replicates. (B) Immunostaining of additional knockout and Foxd4 rescue cell lines differentiated in N2B27  $\pm$  SB43. (C) Quantification of immunostaining of cells at day 5 of differentiation  $\pm$  10 uM SB43 performed by nuclear segmentation and quantitative image analysis. Shown are the mean values for three independent biological replicates. A minimum of 8000 nuclei were scored per experiment. Data for all replicates are shown in Supplementary table 3. Scale bar: 50  $\mu$ m



**Table 1. CRISPR/Cas9 targeting guide RNA and ssODN template sequences.**

Cell line generated	Component	Sequence
E2A-V5	Guide	AGCCGTCACAGCTTCTTCGT
	ssODN	GTCGGGGACCCACAGCTGGCCCTGTCAGCCGCCCCA CCCGGGCCTGGGTGAGGCCCAACAACCCAGCCGGGC ACCTGGGTAAGCCTATCCCTAACCTCTCCTCGGTCT CGATTCTACGTGAGCCGTCACAGCTTCTTCGTTGCAC CAGCGACCACCATATCTCTGCCCGGGGTGCATCAGG ACGGTTCTGGATGAG
E2A <sup>-/-</sup>	Guide	TGCAAACCTGGGTTCCCCCGA
	ssODN	CCTGACTTTTCTCTGTCCCCACAGATGTTCCCGCTAC CAGTGGCCAATGGGAAGAGCCGGCCCGGCTCGCTC GGGTGAGAATTCGGAACCCAGTTTGCAGGCTCAGGT AGGACTGAGGGGCTCTTGGGGGGGTGTTTCATGGGG ACGGTGGCGGAGA
E2A <sup>-/-</sup> HEB <sup>-/-</sup>	Guide	CGGTTGAAATCGTCAGAATT
	ssODN	TATGTAAGTAATACATTTGAATTAATTATCCTGTTTTGT TTTAAGGTATATGCACCATCCGCAAATTCTGACTAGG AATTCCTTAAGTACCCATACGATGTTCCAGATTACGCT GATTTCAACCGTGAATCTCCTAGTTACCCATCTCCCA AGCCAGCAACCAGTATGTTTCGCTAGCACTTTCTTTAT GCAAGGTAAGC

**Table S2. qRT-PCR primers and probes**

Gene	Forward primer	Reverse primer	UPL probe
Cdh1 (E-cadherin)	ATCCTCGCCCTGCTGATT	ACCACCGTTCTCCTCCGTA	18
Cdh2 (N-cadherin)	GCCATCATCGCTATCCTTCT	CCGTTTCATCCATACCACAAA	18
Cripto (Tdgf1)	GTTTGAATTTGGACCCGTTG	GGAAGGCACAACTGGAAAG	93
E2A	GTGGGCTCTGACAAGGAACT	ACAGGTAGCGGGAACATCAT	79
Eomes	ACCGGCACCAAACCTGAGA	AAGCTCAAGAAAGGAAACATGC	9
Fgf5	AAAACCTGGTGCACCCTAGA	CATCACATTCCCGAATTAAGC	29
Foxd4	TGGAGATCAGACGGAAGAAGA	GATCGCTCCAGGCACTTATG	63
Id1	TCCTGCAGCATGTAATCGAC	GGTCCCGACTTCAGACTCC	78
Lefty2	GCCCTCATCGACTCTAGGC	AGCTGCTGCCAGAAGTTCAC	97

Nanog	CCTCCAGCAGATGCAAGAA	GCTTGCACTTCATCCTTTGG	25
Oct4	GTTGGAGAAGGTGGAACCAA	CTCCTTCTGCAGGGCTTTC	95
Pax3	AAAAGGCTAAACACAGCATCG	CAATATCGGAGCCTTCATCTG	110
Sox1	GTGACATCTGCCCCCATC	GAGGCCAGTCTGGTGTCAG	60
SDHA	CAGTTCCACCCCACAGGTA	TCTCCACGACACCCTTCTGT	71
T-brachyury	ACTGGTCTAGCCTCGGAGTG	TTGCTCACAGACCAGAGACTG	27

**Table S3. Immunostaining quantification** (associated with Figure 5)  
**(A)** Cells differentiated in N2B27 **(B)** Cells differentiated in N2B27+SB43

**(A)**

<b>-SB43</b>	<b>% Rep1</b>	<b>% Rep2</b>	<b>% Rep3</b>	<b>Mean %</b>	<b>Standard deviation</b>
<b>Parental</b>					
<b>Oct4-/Sox1+</b>	72.9	89.2	83.8	82.0	8.34
<b>Oct4+/Sox1-</b>	1.8	0.0	1.5	1.1	0.98
<b>Oct4+/Sox1+</b>	6.2	3.1	5.2	4.8	1.60
<b>Oct4-Sox1-</b>	19.1	7.7	9.5	12.1	6.13
<b>E2A-/-</b>					
<b>Oct4-/Sox1+</b>	4.7	7.6	56.5	22.9	29.07
<b>Oct4+/Sox1-</b>	40.3	45.6	3.9	30.0	22.68
<b>Oct4+/Sox1+</b>	3.1	18.2	12.1	11.1	7.59
<b>Oct4-Sox1-</b>	51.9	28.6	27.5	36.0	13.76
<b>E2A-/-HEB-/-</b>					
<b>Oct4-/Sox1+</b>	17.3	6.6	56.0	26.6	28.78
<b>Oct4+/Sox1-</b>	33.4	69.6	2.8	35.3	33.95
<b>Oct4+/Sox1+</b>	3.3	3.3	6.8	4.4	3.91
<b>Oct4-Sox1-</b>	46.1	20.5	34.4	33.7	14.03
<b>E2A-/- +Foxd4 Cl.3</b>					
<b>Oct4-/Sox1+</b>	57.0	61.3	43.5	53.9	9.27
<b>Oct4+/Sox1-</b>	25.5	1.0	0.9	9.1	14.19
<b>Oct4+/Sox1+</b>	11.1	21.4	16.7	16.4	5.15
<b>Oct4-Sox1-</b>	6.4	16.3	38.9	20.5	16.64
<b>E2A-/- +Foxd4 Cl.7</b>					
<b>Oct4-/Sox1+</b>	42.8	45.4	40.0	42.8	2.68
<b>Oct4+/Sox1-</b>	7.1	7.7	7.0	7.3	0.38
<b>Oct4+/Sox1+</b>	8.6	17.7	28.8	18.4	10.12
<b>Oct4-Sox1-</b>	41.5	29.1	24.1	31.5	8.96
<b>E2A-/- +Foxd4 Cl.8</b>					
<b>Oct4-/Sox1+</b>	10.3	0.3	1.9	4.2	5.39
<b>Oct4+/Sox1-</b>	21.5	67.6	64.8	51.3	25.83
<b>Oct4+/Sox1+</b>	5.6	6.4	16.4	9.5	6.00
<b>Oct4-Sox1-</b>	62.5	25.8	16.9	35.1	24.19

(B)

<b>+SB43</b>	<b>% Rep1</b>	<b>% Rep2</b>	<b>% Rep3</b>	<b>Mean %</b>	<b>Standard deviation</b>
<b>Parental</b>					
<b>Oct4-/Sox1+</b>	93.1	99.3	97.7	96.7	3.18
<b>Oct4+/Sox1-</b>	0.0	0.0	0.0	0.0	0.00
<b>Oct4+/Sox1+</b>	0.0	0.7	1.1	0.6	0.56
<b>Oct4-Sox1-</b>	6.9	0.0	1.2	2.7	3.66
<b>E2A-/-</b>					
<b>Oct4-/Sox1+</b>	73.1	65.0	93.9	77.3	14.88
<b>Oct4+/Sox1-</b>	3.8	0.4	0.0	1.4	2.09
<b>Oct4+/Sox1+</b>	3.9	4.9	1.7	3.5	1.63
<b>Oct4-Sox1-</b>	19.2	29.7	4.4	17.8	12.70
<b>E2A-/-HEB-/-</b>					
<b>Oct4-/Sox1+</b>	73.5	63.5	97.8	78.2	17.19
<b>Oct4+/Sox1-</b>	0.0	0.2	0.0	0.1	0.09
<b>Oct4+/Sox1+</b>	1.7	1.9	1.3	1.7	1.19
<b>Oct4-Sox1-</b>	24.7	34.4	0.9	20.0	16.92
<b>E2A-/- +Foxd4 Cl.3</b>					
<b>Oct4-/Sox1+</b>	82.8	85.1	96.7	88.2	7.44
<b>Oct4+/Sox1-</b>	0.0	0.0	0.0	0.0	0.00
<b>Oct4+/Sox1+</b>	3.6	1.5	2.6	2.6	1.04
<b>Oct4-Sox1-</b>	13.5	13.4	0.6	9.2	7.40
<b>E2A-/- +Foxd4 Cl.7</b>					
<b>Oct4-/Sox1+</b>	85.1	94.4	97.0	92.2	6.27
<b>Oct4+/Sox1-</b>	0.0	0.0	0.0	0.0	0.00
<b>Oct4+/Sox1+</b>	1.8	0.6	3.0	1.8	1.16
<b>Oct4-Sox1-</b>	13.1	5.0	0.0	6.0	6.60
<b>E2A-/- +Foxd4 Cl.8</b>					
<b>Oct4-/Sox1+</b>	76.3	95.6	97.0	89.6	11.58
<b>Oct4+/Sox1-</b>	0.0	0.0	0.0	0.0	0.00
<b>Oct4+/Sox1+</b>	0.0	2.3	3.0	1.8	1.58
<b>Oct4-Sox1-</b>	23.7	2.1	0.0	8.6	13.14